9-26-2013 12:00 AM

PEGylation as a novel tool to investigate the topology of Escherichia coli WecA, a membrane enzyme involved in lipopolysaccharide O antigen initiation

Stéphanie L. Lamothe
The University of Western Ontario

Supervisor
Dr. Miguel A. Valvano
The University of Western Ontario

Graduate Program in Microbiology and Immunology
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

PEGylation as a novel tool to investigate the topology of *Escherichia coli* WecA, a membrane enzyme involved in lipopolysaccharide O antigen initiation

(PEGylation of *E. coli* WecA)
(Monograph)

By

Stéphanie Lamothe

Graduate Program in Microbiology and Immunology

Submitted in partial fulfillment of the
requirement for the degree of
Master of Science

School of Graduate and Post-Doctoral Studies
The University *of* Western Ontario
London, Ontario, Canada

Abstract

O-antigen, the most surface exposed moiety of bacterial lipopolysaccharide (LPS), plays several roles in pathogenicity. The biosynthesis of O-antigen starts by the formation of a phosphoanhydride bond linking a sugar phosphate with a membrane isoprenoid lipid phosphate. Two distinct families of integral membrane proteins catalyze this reaction. The protein WecA is the prototypic member of one of these families, termed the polyisoprenyl-phosphate *N*-acetylaminosugar-1-phosphate transferase (PNPT) family. Because the donor nucleotide sugar is only available in the cytosol, cytosolic exposed regions of WecA are expected to be critical for the catalytic activity of the enzyme. Therefore, elucidating an accurate topological map of WecA is essential to understand its function. While a topological model has been determined with some level of accuracy in certain regions of WecA, various protein-protein, protein-lipid, and protein-aqueous interfaces are not precisely mapped and we have also found inconsistencies between *in silico* prediction models and experimental data. We hypothesize that the borders between the transmembrane domains (TMs) and the cytosolic loops of WecA are critical for its function. In particular, this thesis focuses on cytoloops 1 and 4 (including the universally conserved VFMGD motif), and TMs IV and V (a putative interaction site for the lipid substrate). The locations of these border residues were identified by the topological accessibility of specific amino acids using the substituted cysteine accessibility method (SCAM) combined with PEGylation. This approach offers several advantages over other classical SCAM methods, especially by avoiding the need to purify the labeled protein after sulfhydryl chemistry. PEGylation involves the covalent mass labeling of accessible sulfhydryl groups with the large thiol reagent methoxy-polyethylene glycol-maleimide (PEG-Mal, 5 kDa), and detection using a gel shift assay. Orientation is differentiated by treatment of EDTA-permeablized cells compared to treatment in membrane preparations. This thesis describes the adaptation and use of PEGylation for the topological analysis of WecA. The results provide further refinement of the WecA topological map, and the method can be extended to other integral membrane proteins involved in O-antigen assembly.

Keywords: Lipopolysaccharide, O-antigen initiation, WecA, PEGylation, PEG-Mal

Acknowledgements

Foremost, I would like to thank Dr. Miguel Valvano for his unparalled support and encouragement throughout this project. His continued interest and enthusiasm for research and advancing scientific thought is admirable. I would also like to thank the many members of the Valvano lab, past and present, whom I have had the pleasure of learning from over the last two years. I would especially like to thank my advisory committee members, Dr. Creuzenet and Dr. McCormick, for advice throughout this project. Finally, I would like to thank my friends and family who have been more than gracious with their unwavering encouragement, comfort, and support over the past two years.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

μL: Microlitre (s)

A: Alanine

Ab: Antibody

B. cenocepacia: *Burkholderia cenocepacia*

BSA: Bovine serum albumin

C: Cysteine

CPS: Capsular polysaccharide or capsule

Cytoloop (s): Cytoplasmic loop (s)

D: Aspartic acid

ddH_2O: Double-distilled water

Dol-P: Dolichyl phosphate

ECA: Enterobacterial common antigen

E. coli: *Escherichia coli*

EDTA: Ethylenediaminetetracetic acid

F: Phenylalanine

G: Glycine

Gal: D-galactose

Glc: D-glucose

GalNAc: N-acetylgalactosamine

GlcNAc: N-acetylglucosamine

H: Histidine

HIGH: Histidine-isoleucine-glycine-histidine motif

HIHH: Histidine-isoleucine-histidine-histidine motif

h: Hour (s)

I: Isoleucine

K: Lysine

kDa: Kilo Dalton (s)

Kdo: α-3-*deoxy*-D-*manno*-oct-2-ulosonic acid

L: Leucine

L: Litre (s)

LacZ: β-galactosidase

LB: Luria-Bertani medium

LPS: Lipopolysaccharide

M: Methionine

min: Minute (s)

mL: Mililitre (s)

MraY: UDP-MurNAc-pentapeptide:undecaprenylphosphate MurNAc-pentapeptide-1-P transferase

MurNAc: N-acetylmuramoyl

MTSET: 2-(trimethylammonium)ethyl methanethiosulfonate

N: Asparagine

OAg: O-antigen

OD: Optical density (nm)

OS: Core oligosaccharide

O unit: O-antigen subunit

PEG: Poly(ethylene) glycol

PEG-Mal: monoMethoxy-Poly(ethylene) glycol-maleimide

PhoA: Alkaline-phosphatase

PHPT: Polyisoprenyl-phosphate hexose-1-phosphate transferases

PIRS: Polyisoprenyl recognition sequence

PNPT: Polyisoprenyl-phosphate *N*-acetylaminosugar-1-phosphate transferases

Q: Glutamine

R: Arginine

S: Serine

SCAM: Scanning Cysteine Accessibility Method

SDS: Sodium dodecyl sulfate

SDS-PAGE: Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

sec: Second (s)

SOB: Super optimal broth

T: Threonine

TBS: Tris-buffered saline

TM: Transmembrane domain

TMHMM: Transmembrane Hidden Markov Model

UDP: Uridine diphosphate

Und-P: Undecaprenyl phosphate

UDP-GalNAc: UDP-N-acetylgalactosamine

UDP-GlcNAc: UDP-N-acetylglucosamine

UDP-FucNAc: UDP-N-acetyl-D-fucosamine

UDP-QuiNAc: UDP-N-acetyl-D-quinovosamine

UMP: Uridine monophosphate

V: Valine

V: Volts

W: Tryptophan

WaaL: O-antigen ligase

WbaP: UDP-Gal:undecaprenylphosphate Gal-1-P transferase

WbcO: UDP-FucNAc:undecaprenylphosphate FucNAc-1-P transferase and/or UDP-QuiNAc:undecaprenylphosphate QuiNAc-1-P transferase

WbpL: UDP-FucNAc:undecaprenylphosphate FucNAc-1-P transferase and/or UDP-QuiNAc:undecaprenylphosphate QuiNAc-1-P transferase

(v/v): Volume by volume

wecA: UDP-GlcNAc:undecapreylphosphate GlcNAc-1-P transferase gene

WecA: UDP-GlcNAc:undecapreylphosphate GlcNAc-1-P transferase

(w/v): Weight by volume

Wzx: O-antigen "flippase" translocase

Wzy: O-antigen polymerase

Wzz: O-antigen chain length regulator

X-Gal: 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside

Y: Tyrosine

List of Appendices

Chapter 1

Introduction

Introduction

1.1 The bacterial cell envelope

The cell envelope of Gram-negative bacteria prevents harmful compounds from entering the bacterial cell, maintains the cell's rigidity and shape, and allows entry of nutrients [1, 2]. The cell envelope consists of an inner membrane, a peptidoglycan layer, and an outer membrane (Figure 1). The inner membrane is a phospholipid bilayer composed primarily of phospholipids and integral membrane proteins. Phosphatidylethanolamine and diphosphatidylethanolamine are the most abundant lipids in the inner membrane. The enzymes responsible for the synthesis of peptidoglycan and select outer membrane components are major proteins located in the inner membrane, in addition to secretion systems, efflux pumps, membrane sensor kinases, and components of the respiratory chain [3].

Glycan chains of N-acetylglucosamine and N-acetylmuramic acid cross-linked by a short polypeptide form a net-like polymeric structure known as cell wall peptidoglycan or murein. The chemistry of cross-links and the peptide composition vary between bacterial species (reviewed in [4]). Essential for cell survival, the peptidoglycan cell wall maintains cell integrity by withstanding internal turgor (osmotic) pressure. Also, the peptidoglycan layer acts as a scaffold for other cell envelope components, such as the peptidoglycan-associated lipoprotein that anchors the outer membrane to the peptidoglycan [2].

The outer membrane is an asymmetric bilayer composed of phospholipids, lipoproteins, polysaccharides such as capsular polysaccharides (CPS or capsule), enterobacterial common antigen (ECA), lipopolysaccharide (LPS), and proteins, and is

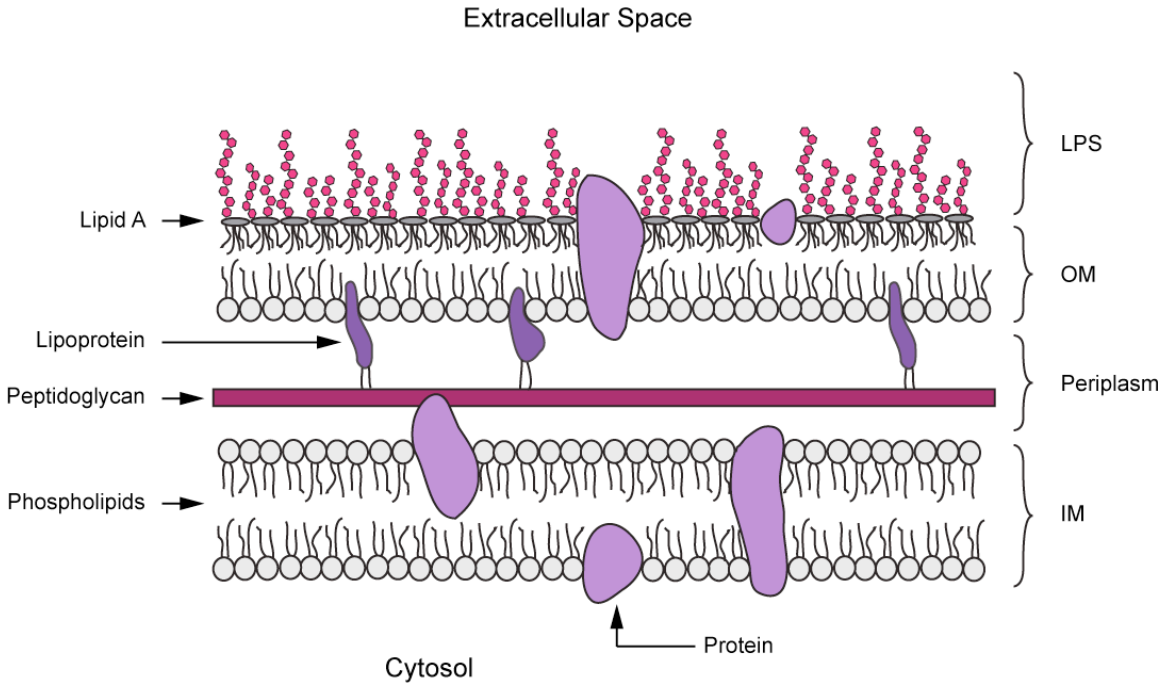


Figure 1. Schematic representation of the Gram-negative bacterial cell envelope (Modified from [5]). The cell envelope consists of 3 layers: the inner membrane (IM), the periplasmic space, including the peptidoglycan, and the outer membrane (OM). Proteins are found in all 3 layers and surface polysaccharides like LPS protrude out of the outer membrane into the extracellular space.

selectively permeable by preventing the diffusion of macromolecules, neutral and anionic substances and hydrophobic molecules [6-8]. The outer membrane confers Gram-negative bacteria resistance to various bactericidal compounds such as bile salts, hydrophobic antibiotics, and detergents [2, 9, 10]. These properties of the outer membrane also help the bacteria to evade host defenses such as phagocytosis and complement killing [11, 12]. A defining feature of the outer membrane bilayer is its asymmetry; the inner leaflet has a similar lipid composition as the inner membrane while the outer leaflet has a low concentration of phospholipids with LPS being the dominant structural component [2, 9, 10].

1.2 Lipopolysaccharides

1.2.1 Structure of LPS

LPS is a heat-stable and negatively charged surface glycolipid that has 3 structural components: lipid A, core oligosaccharide (OS), and O-antigen (OAg). LPS is anchored to the outer leaflet of the outer membrane by the hydrophobic lipid A portion, which is also the most conserved LPS moiety (reviewed in [13]; Figure 2).

The next most conserved structural component of LPS is the core OS, consisting of 8 to 12 sugar residues in total, which can be further subdivided into an inner core and outer core. Rare sugars like Kdo and L- or D-*glycero*-D-*manno*-heptose are characteristic of the inner core OS (reviewed in [14]), and modifications, like phosphorylation, are common to the inner core heptoses. D-glucose (Glc), D-galactose (Gal), D-glucosamine (GlcN), N-acetylglucosamine (GlcNAc), and N-acetylgalactosamine (GalNAc) are common hexoses found in the outer core OS.

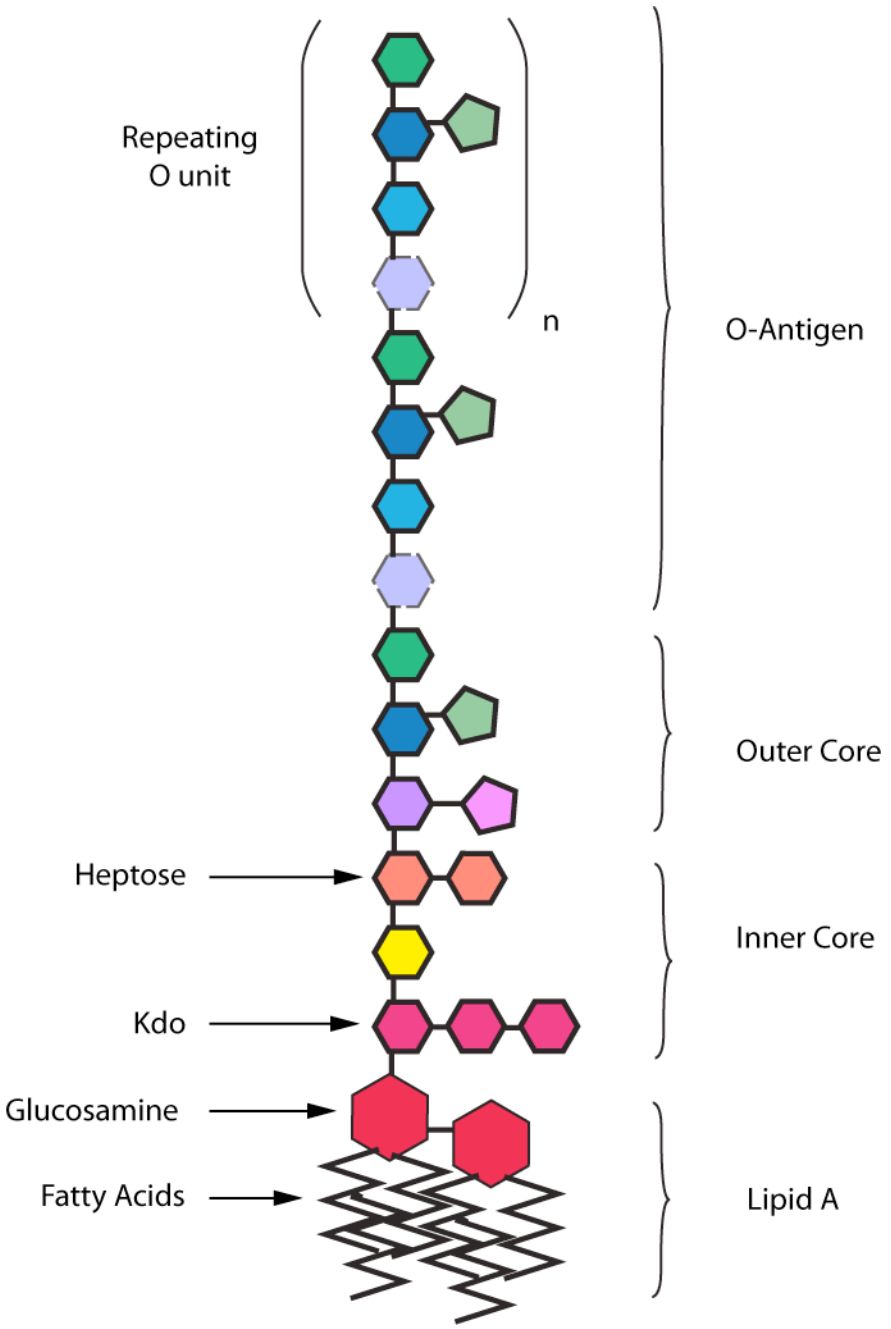


Figure 2. General structure of lipopolysaccharide (Modified from [5]). The three main structural components of LPS are: the lipid A, the core oligosaccharide, including the inner and outer core, and the O-antigen. Kdo = α-3-*deoxy*-D-*manno*-oct-2-ulosonic acid; n= number of O units.

The outer core OS is the site of attachment for the OAg, which is the most surface exposed LPS structural component.

The OAg is also the most heterogeneous structural component of LPS. In *E. coli* alone, 170 serologically different OAg molecules have been identified [15]. OAg is a polymer of polysaccharide subunits or O units, which are composed of unique monosaccharides (reviewed in [16]) and can be linear or branched. Due to the variation in types of sugars, their sequence in the subunit, and the type of linkages between sugars, there are numerous possible O unit structures. Therefore, OAgs are highly variable across species and even strains. Additional variability exists through modifications such as acetylation or glycosylation. The variation in the OAg structure accounts for a wide range of serotypes in Gram-negative bacterial species.

1.2.2 LPS biosynthesis

The lipid A-core and OAg are synthesized in separate biosynthetic pathways. The core OS is assembled onto preformed lipid A by sequential glycosyltransfer reactions while OAg is assembled onto undecaprenyl-phosphate (Und-P) forming an Und-PP-linked saccharide. Both synthesis pathways are initiated at the cytoplasmic face of the inner membrane and require a translocation step to the periplasmic face of the inner membrane. MsbA, an inner membrane ATP Binding Cassette (ABC)-transporter, is responsible for the transport of the lipid A-core across the inner membrane (reviewed in [13]) while the OAg biosynthesis, initiated by WecA, will be discussed in section 1.3. Once both molecules are translocated across the inner membrane to the periplasmic face, they become ligated together with the release of Und-PP. Und-PP is then

dephosphorylated into Und-P by a poorly characterized but conserved pathway. Finally, the complete LPS molecule is transported through the periplasmic space and then onto the outer leaflet of the outer membrane by the Lpt pathway [17].

1.3 Biosynthesis and assembly of O-antigen

OAg biosynthesis starts at the cytosolic face of the inner membrane with the formation of an Und-PP-linked intermediate. All cells produce polyisoprenoids, which are lipid carrier intermediates for the biosynthesis of complex carbohydrate structures. Polyisoprenol-phosphate-linked saccharides are needed in the early stages of protein glycosylation in eukaryotes and prokaryotes, as well as for bacterial cell wall peptidogylcan and surface polysaccharide synthesis. Nucleotide sugars, available as soluble molecules in the cytosol, donate carbohydrates for the synthesis of the saccharide moiety. In contrast, the phosphoisoprenol acceptor is embedded within the lipid membrane bilayer. After assembly, the phosphoisoprenol-linked saccharide molecules must cross the lipid bilayer for further processing. Therefore, the transmembrane movement of phosphoisoprenol-linked saccharides represents a conserved, obligatory step of significant biological importance in all types of cells [17-21].

Depending on the specific type of assembly system, the Und-PP-OAg is subsequently elongated, translocated across the membrane, and polymerized, ultimately resulting in an Und-PP-linked polymeric OAg localized to the periplasmic side of the inner membrane. Finally, the OAg glycan moiety becomes ligated onto the lipid A-core OS resulting in the formation of a complete LPS molecule while the Und-PP polyisoprenoid carrier is dephosphorylated into Und-P by several membrane-associated

phosphatases and recycled to reinitiate OAg biosynthesis [17-20, 22, 23]. Two main pathways are responsible for the assembly of OAg: the Wzy/Wzx-dependent pathway and various types of ABC-transporter-dependent pathways.

1.3.1 Wzy/Wzx-dependent O-antigen biosynthesis

The Wzy/Wzx pathway involves the synthesis of O-units by the successive addition of monosaccharides at the non-reducing end of the molecule. This pathway is responsible for the synthesis of the majority of OAgs, and especially heteropolymeric OAgs. Wzx is an inner membrane protein with 12 predicted TMs that translocates the O-units across the inner membrane. Wzx proteins share little primary sequence similarity such that they can be used as genetic markers for distinguishing among specific OAgs. Wzy polymerizes the translocated O-units. Polymerization proceeds sequentially from the reducing end of the growing polymer to the non-reducing end of Und-PP-linked units. Wzy proteins are also inner membrane proteins with 11-13 predicted TMs and also share little sequence similarity. Unlike Wzx, Wzy proteins from different O-types are not interchangeable and display specificity for the O-unit. Several Wzy proteins were examined with topology prediction programs and it appears that Wzy proteins possess a large periplasmic loop that may be important in the recognition of the OAg subunit.

The third component to the Wzy/Wzx-dependent pathway is the chain-length regulating protein Wzz. These proteins are in the inner membrane, and all have a periplasmic loop with a predicted coiled-coil structure flanked by two TMs. Wzz controls the degree of polymerization by Wzy and appears not to be specific for a given O-repeat subunit structure. It is thought these three proteins work together as a functional complex

(Figure 3). Once the OAg molecule is complete, the integral membrane ligase WaaL transfers the complete OAg from Und-P ‘en bloc’ onto the pre-formed lipid A-core to form a covalent, glycosyl linkage between the two molecules, resulting in a complete LPS molecule which is exported to the outer surface of the outer membrane [24-35].

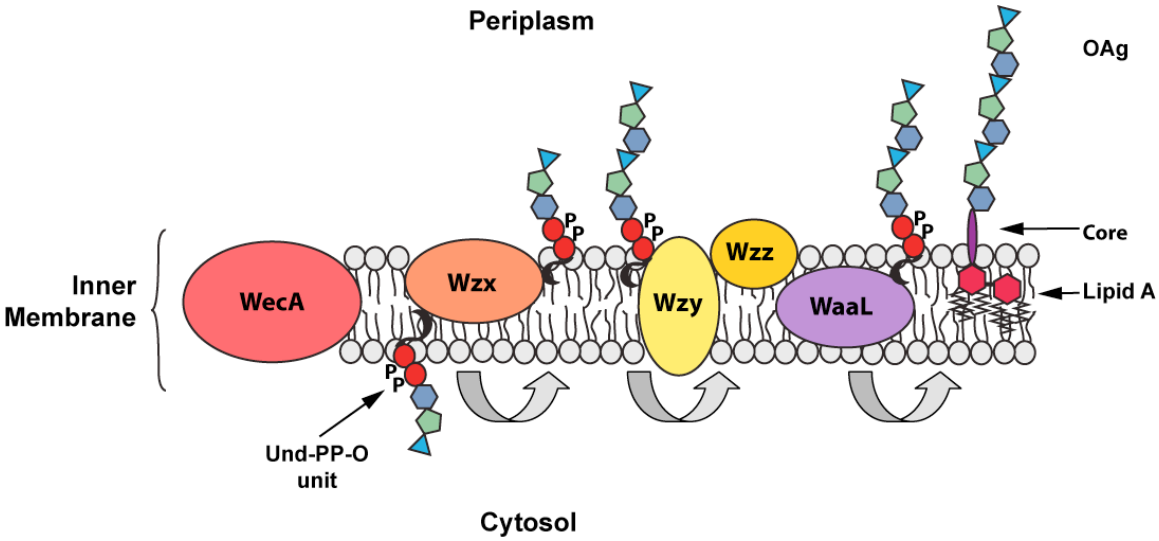


Figure 3. Wzy/Wzx-dependent O-antigen biosynthesis pathway in *E. coli* (Modified from [5]). Initiation of OAg biosynthesis occurs at the cytosolic face of the inner membrane. WecA catalyzes the phosphoanhydride bond formation between GlcNAc-1P and Und-P to yield Und-PP-GlcNAc. Membrane associated glycosyltransferases add additional sugars onto the Und-PP-GlcNAc intermediate forming O units. Wzx translocates the O unit across the inner membrane to the periplasmic face where Wzy polymerizes the O units together. Wzz regulates the degree of polymerization. Upon completion of polymerization, WaaL transfers the OAg from Und-PP to the preformed lipid A-core, and covalently ligates the LPS molecule together.

1.3.2 ABC transporter-dependent O-antigen biosynthesis

The ATP-binding cassette (ABC) transporter pathway translocates the complete OAg across the inner membrane in a process that requires ATP hydrolysis. This pathway operates primarily in the synthesis of linear, homopolymeric OAg molecules. ABC transporters are also involved in the export of lipid-linked glycans for the assembly of glycoproteins, teichoic acids, and CPSs. Seven different classes can be identified that represent ABC transporter-dependent pathways. Six of these classes consist of independent pairs of TM polypeptides and polypeptides containing the nucleotide-binding domains (NBDs) [17, 36].

Elongation of the entire OAg occurs at the cytoplasmic face of the inner membrane by glycosyltransferases that add sugars to the Und-PP-sugar (Figure 4). Unlike Wzy-dependent synthesis, the GlcNAc residue is transferred to lipid A-core OS during ligation only once per chain, therefore it is not found within the O-repeat unit structure itself. The O-specific polysaccharides formed by this pathway display strain-specific chain length (modal) distributions despite that the ABC-transporter dependent pathway does not involve a Wzz protein. The chain-terminating WbdD protein modifies the nonreducing end of these polymers, causing termination of polymerization, therefore controlling the length of these polymers [13, 17, 36].

ABC-2 transporters contain a TMD component, Wzm, an inner membrane protein with an average of six TMs, and a hydrophilic NBD protein, designated as Wzt. Wzm homologues display little primary sequence similarity, while Wzt homologues are much more highly conserved, especially in the NBD domain. Contrarily, Wzm proteins are interchangeable between different OAgs, while Wzt proteins are not. Again, once

translocated, WaaL mediates ligation of the OAg onto lipid A-core at the periplasmic face of the inner membrane [13, 37].

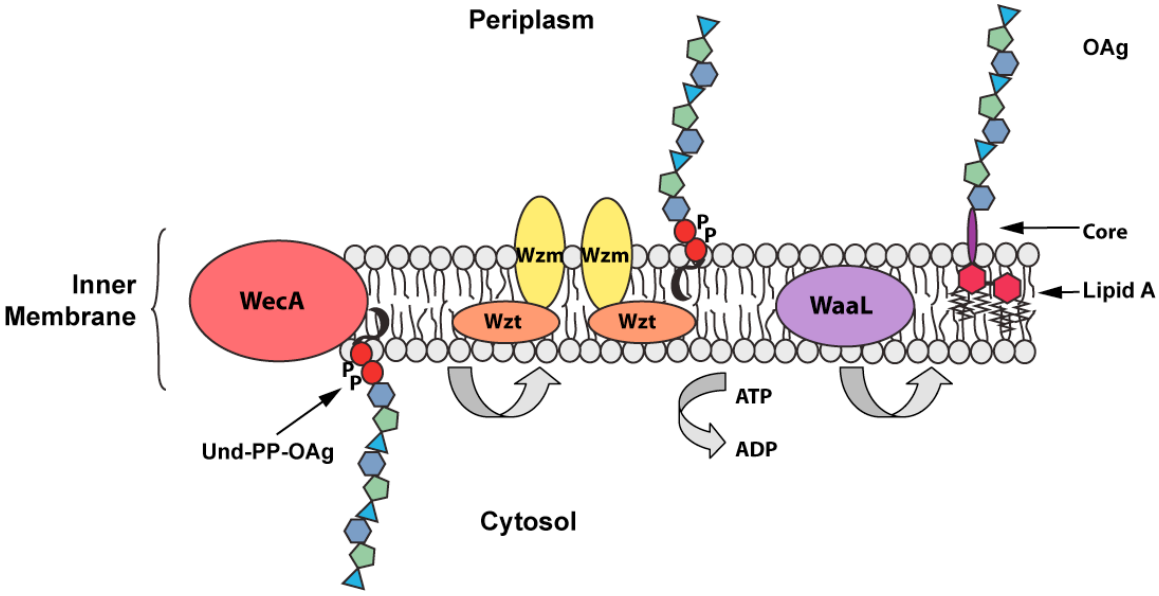


Figure 4. ABC transporter-dependent O-antigen biosynthesis pathway in *E. coli* (Modified from [5]). Initiation of OAg biosynthesis occurs at the cytosolic face of the inner membrane to form the isoprenoid linked OAg intermediate. Soluble glycosyltransferases add additional sugars onto the intermediate to complete OAg. The ABC transporter, composed of Wzt and Wzm, translocates the complete OAg across in the inner membrane to the periplasmic face. WaaL covalently ligates the OAg from Und-P to the preformed lipid A-core.

1.3.3 Synthase-dependent OAg biosynthesis

The O:54 antigen of *S. enterica* serovar Borreze is the only known example of the synthase-dependent O-polysaccharide. The O:54-specific polysaccharide is a linear homopolymer of *N*-acetylmannosamine (ManNAc). Synthases are inner membrane proteins that catalyze a vectorial polymerization reaction by a processive mechanism resulting in the extension of the polysaccharide chain with the simultaneous extrusion of the nascent polymer across the inner membrane. The exact mechanism of elongation, translocation, and chain termination are still unknown [17, 36].

1.4 Initiating enzymes

Two major families of enzymes catalyze the initiation reaction of OAg biosynthesis: Polyisoprenyl-phosphate *N*-acetylaminosugar-1-phosphate transferases (PNPTs) and Polyisoprenyl-phosphate hexose-1-phosphate transferases (PHPTs) [36]. See figure 5 for a schematic topological comparison of the two major families.

1.4.1 PHPT family

All PHPT transferases are inner membrane proteins found only in prokaryotes, and share a high level of primary sequence similarity. The prototypic enzyme of this family is WbaP, which initiates OAg biosynthesis in *S. enterica* by transferring galactose-1-phosphate from UDP-galactose (UDP-Gal) to Und-P, yielding Und-P-P-Gal [38-41]. WbaP has three domains with five predicted TM helices. The hydrophobic N-terminal domain spanning four TMs has no known function but contributes to the overall stability of the protein in the membrane [40]. The large, central periplasmic loop between TM 4

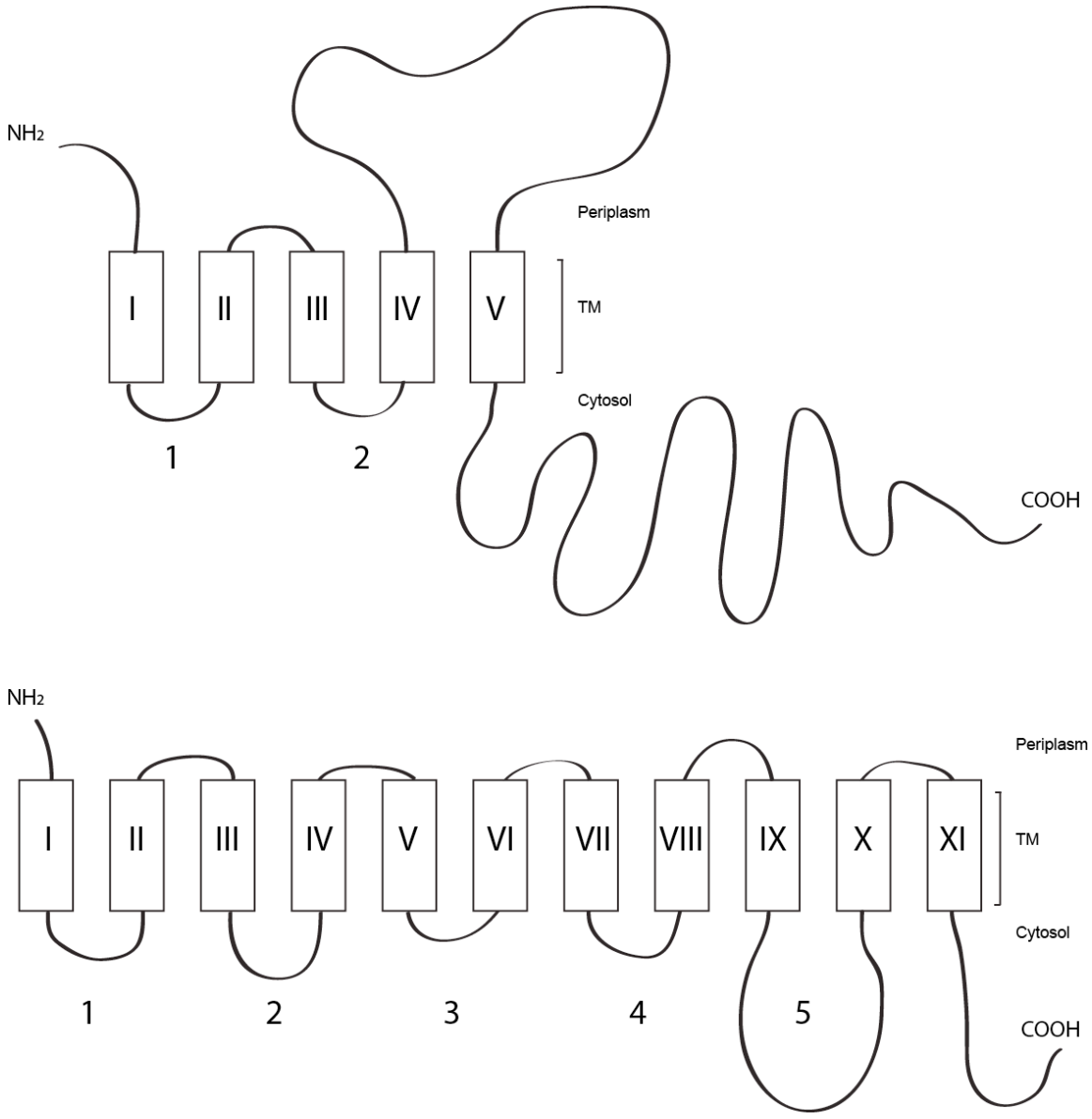


Figure 5. Schematic representation of the topologies of the PHPT and the PNPT prototypic enzymes. Roman numerals indicate the TM helices. Numbers indicate the cytosolic loops. A) Topological representation of WbaP from *S. enterica* serovar Typhi. The topology was predicted using the TMHMM, MEMSAT, TOPPRED, and PHD programs. B) Topological representation of WecA from *E. coli*. The model was originally derived using the TMHMM program and the TM boundaries were experimentally refined by SCAM [20].

and TM 5 is implicated in modulating chain length distribution of OAg [39, 42]. Most importantly, the C-terminal domain, including TM5 and the tail, is sufficient for *in vitro* and *in vivo* enzymatic activity, and its localization was confirmed by trypsin cleavage and GFP reporter experiments [39, 42]. Other members of the PHPT family initiate the synthesis of various types of glycans such as: colonic acid in *E. coli* K-12 [39, 42, 43], K30 capsule in *E. coli* 09:K30 [42, 44, 45], type 2 capsule in *Streptococcus pneumoniae* [39, 46, 47], and glycans for S-layer protein glycosylation in *Geobacillus stearothermophilus* [39, 48] and general protein glycosylation [36].

1.4.2 PNPT family

The PNPT family of proteins has prokaryote and eukaryote members. PNPT enzymes catalyze the transfer of N-acetylaminosugar-1-phosphate from the donor, UDP-N-acetyl-aminosugar to the lipid acceptor. Und-P is the only lipid acceptor found in bacteria, whereas the lipid acceptor in eukaryotes and Archaea is dolichylphosphate (Dol-P) [49]. Dol-P and Und-P both contain an α-isoprene unit. This is the phosphorylated end of the molecule that participates in the phosphoanhydride bond formation with the N-acetylaminosugar-1-P. Und-P contains 11 isoprene units, all of which are fully unsaturated, while Dol-P can be made of 15-19 isoprene units that have a saturated α-isoprene [50]. Eukaryotic and bacterial evolutionary divergence may be reflected by their ability to discriminate their lipid substrate (Figure 6).

Eukaryotic PNPTs are involved in asparagine N-linked protein glycosylation, transferring GlcNAc-1-phosphate from UDP-GlcNAc to Dol-P forming the product

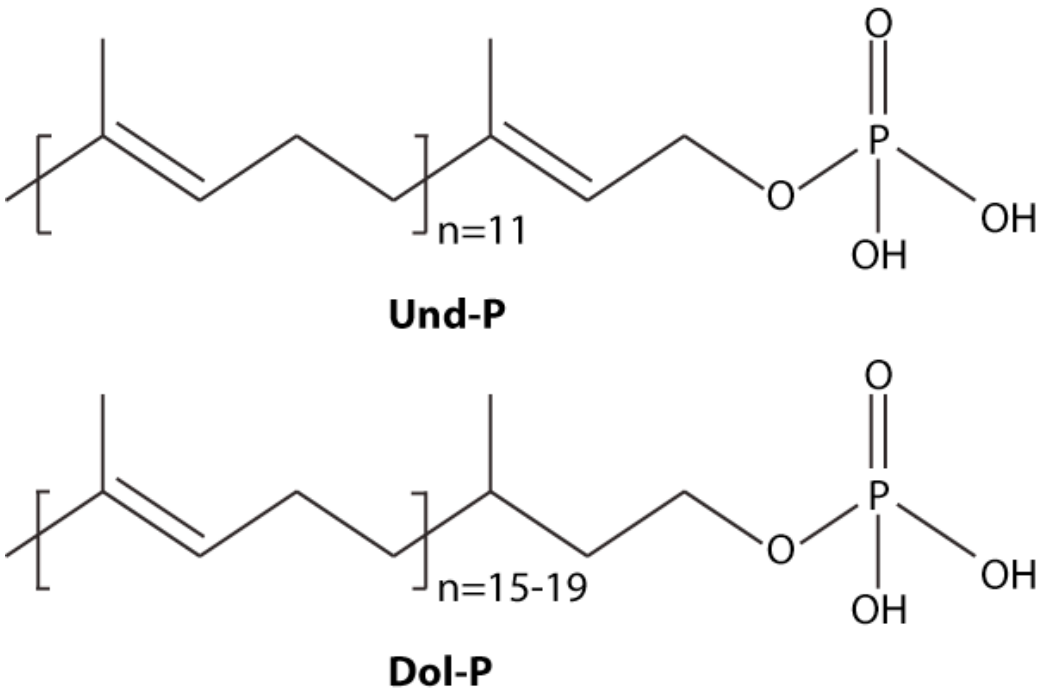


Figure 6. Lipid carriers utilized by the PNPT family of enzymes (Modified from [17]). Structure of isoprenoid lipids. A) Und-P. B) Dol-P. Und-P has 11 isoprene units and Dol-P has 15-19 isoprene units. The α isoprene unit is the one linked to the phosphate mocule. n = number of total isoprene units.

Dol-PP-GlcNAc [51-53]. The bacterial PNPT WecA, the prototype for the PNPT family, is a UDP-GlcNAc:UndP GlcNAc-1-phosphate transferase. This reaction is involved in initiation of OAg and ECA synthesis in enteric bacteria [26, 27, 54-58]. Contrary to other PNPT enzymes, WecA has relaxed specificity and can also use UDP-GalNAc [54, 59]. Other bacterial PNPTs include WbpL, WbcO, and MraY [59-62]. WbpL and WbcO are thought to be specific for UDP-N-acetyl-D-fucosamine (UDP-FucNAc) and/or UDP-N-acetyl-D-quinovosamine (UDP-N-QuiNAc). MraY proteins initiate peptidoglycan synthesis by catalyzing the transfer of N-acetylmuramoyl-pentapeptide 1-phosphate (MurNAc-pentapeptide-1P) from UDP-MurNAc-pentapeptide to Und-P forming Und-PP-MurNAc-pentapeptide.

Prokaryotic and eukaryotic PNPTs are polytopic inner membrane proteins sharing six discrete, conserved domains found in bacterial and eukaryotic homologues [22, 23, 53, 63-66]. Due to sequence similarity and location at the transmembrane:cytosol interface, these domains were proposed to be involved in catalysis or substrate binding [53]. The PNPT family members have five cytosolic loops (cytoloops). Conserved regions in cytoloops 2, 3, and 4 (including the highly conserved VFMGD motif) are found in all bacterial PNPT members and likely contribute to form the catalytic site of these enzymes [51, 53, 65, 67]. A high degree of variability was noted in the large cytoloop 5, which is proposed to be involved in nucleotide sugar specificity (Figure 7) [51, 65].

1.5 Structure-function of WecA

1.5.1 Topology of WecA

WecA is a 367 amino-acid bacterial sugar-phosphate transferase with a molecular mass of 40.9 kDa. The WecA topology was initially predicted *in silico* using the transmembrane hidden Markov model (TMHMM) and compared to known topologies of MraY and other PNPT family proteins [20, 22, 51, 67, 68]. Reporter fusions and the substituted cysteine accessibility method (SCAM) were utilized to refine the predicted model and confirm that WecA has 11 transmembrane helices (TMs), 5 cytoplasmic loops, 5 periplasmic loops, and a cytosolic C-terminus (Figure 7) [20, 51].

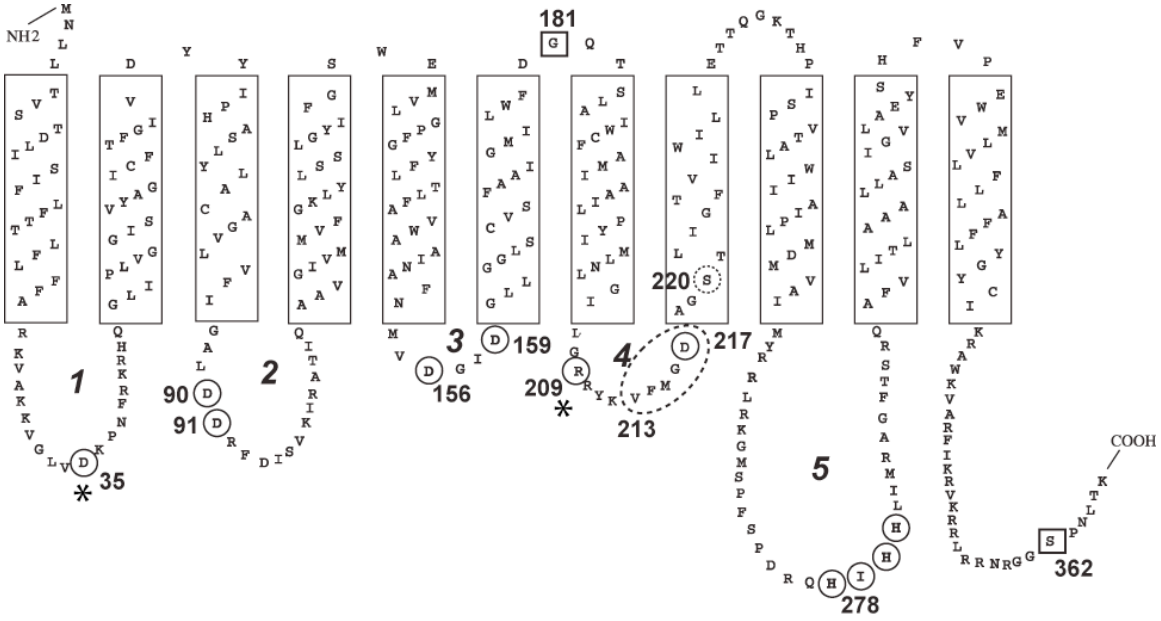


Figure 7. Predicted topology of *E. coli* MV501 WecA (Modified from [51]). The N-terminus is in the periplasm, the 5 cytoloops are indicated by numbers 1-5, and the C-terminus is in the cytoplasm. The dotted-line circle indicates the updated location of the highly conserved VFMGD motif (based on biotinylation of aspartic acid 217). The asterisks indicate the controversial residues D35 and R209.Other critical residues are indicated by circles with white backgrounds while the squares indicate the periplasmic and cytoplasmic controls.

1.5.2 Regions of WecA essential for catalysis

Numerous attempts have been made to define regions and individual amino acids in WecA that are functionally significant. TMs I through III are essential for function but not for membrane insertion [69]. TM XI is essential for stability of the protein in the membrane while the C-terminal tail of WecA is dispensable for function [22]. The highly conserved arginine (R) at position 265, and the histidine-isoleucine-histidine-histidine (HIHH) motif at positions 279-282 within cytoloop 5 are thought to play a role in UDP-GlcNAc recognition [22, 65]. The HIHH sequence is reminiscent of the histidine-isoleucine-glycine-histidine (HIGH) motif conserved in the active site of nucleotide phosphate binding proteins, such that amino acid replacements of R_{265} and the $HIHH_{279-282}$ motif resulted in loss of OAg production and binding of the transition state analogue, tunicamycin [22, 70, 71]. Replacement of H_{279} with glycine (G) demonstrated that this residue is critical for enzyme function [71].

Aspartic acid (D) residues at positions 90 and 91 in cytoloop 2, and 156 and 159 in cytoloop 3 are also conserved in the PNPT family but are functionally distinct [22, 23]. The conserved aspartic acids in cytoloop 2 of WecA were proposed to bind the cofactor Mg^{2+}, while D156 and D159 were proposed to be involved in UDP-GlcNAc binding or modification. Only D156 is essential for catalysis. Amino acid substitutions in all of the conserved aspartic acids cause defects in OAg initiation, supporting these hypotheses [20, 23].

Eukaryotic and prokaryotic PNPT family members share another highly conserved region, the VFMGD motif (Figure 8). This motif in MraY was proposed to contain a catalytic nucleophile, D267 [72], while experiments in our laboratory suggest

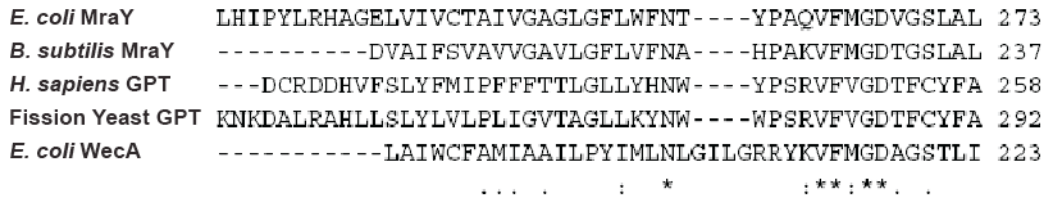


Figure 8 (Modified from [51]). ClustalW alignment of the highly conserved VFMGD motif among PNPT family members. Alignment shows protein sequences from *E. coli* MraY, *B. subtilis* MraY, *H. sapiens* GPT, Fission Yeast GPT, and *E. coliI* WecA. An asterisk (*) indicates completely conserved amino acids, and a colon (:) indicates partially conserved amino acids.

the comparable motif and D217 in WecA does not [51]. The size of the amino acid side chains and the polarity at the D217 position are important for enzymatic activity, as were the replacements of the surrounding residues with alanine. Therefore this highly conserved motif is proposed to define a region in PNPT proteins that contributes to the active site [51].

1.6 Methods for predicting and determining inner membrane protein topologies

Topological methods provide insight into essential amino acids, organization, and structure of inner membrane proteins. There are several methods to predict and experimentally probe the topology of membrane proteins.

1.6.1 *In silico* analysis

The precise identification of TMs is important for the functional annotation and the functional analysis when studying any membrane protein. A major obstacle when studying membrane proteins is the difficulty in experimental determination of their three-dimensional structures, especially if no prior topology prediction studies have been performed. Experimental studies are therefore complemented, and generally initiated, by computational studies using neural networks, Hidden Markov models, and support vector machines [73]. Most topology predictions methods, such as the Transmembrane Hidden Markov Model (TMHMM), are based on processing multiple-sequence alignment data, and several assumptions such that hydrophobic regions are potential TMs, coiled-coil regions are potential protein-interaction domains, and hydrophilic stretches are potential loops [73, 74]. TM predictions are possible since most TM α-helices contain long

stretches of 17-25 non-polar hydrophobic residues. These hydrophobic regions cross the lipid membrane many times, and contain connecting loops with more polar residues. The orientation of the helices, whether inward or outward, is also important to the overall topology of the protein. This is highlighted by the well-established "positive-inside" rule that predicts that a hydrophobic region followed by positively charged residues, such as arginine and lysine, may indicate an inward orientation of the helix [73-76]. The TMHMM program takes advantage of these postulates and is therefore reasonably accurate, providing near 80% confidence based on analyses of model membrane proteins [73, 75]. *In silico* methods for predicting membrane protein topology provide a general topological model of the protein of interest, especially when different computer programs afford similar predictions [68, 73, 76]. However, non-canonical residues in TMs, like positively or negatively charged amino acids can lead to wrong predictions.

1.6.2 Reporter fusions

Topological arrangement of proteins within the bacterial cytoplasmic membrane has been successfully studied using gene fusions techniques. The most widely used reporters are β-galactosidase (LacZ), and alkaline phosphatase (PhoA).

The periplasmic alkaline phosphatase PhoA and the cytoplasmic β-galactosidase LacZ are reporter proteins that have reciprocal activity in different cellular compartments; therefore, these enzymes are used as sensors for the subcellular locations of segments of integral membrane proteins (Figure 9). PhoA is only enzymatically active when it is translocated across the inner membrane into the periplasmic space. This is visualized when the bacteria expressing enzymatically active PhoA appear blue when

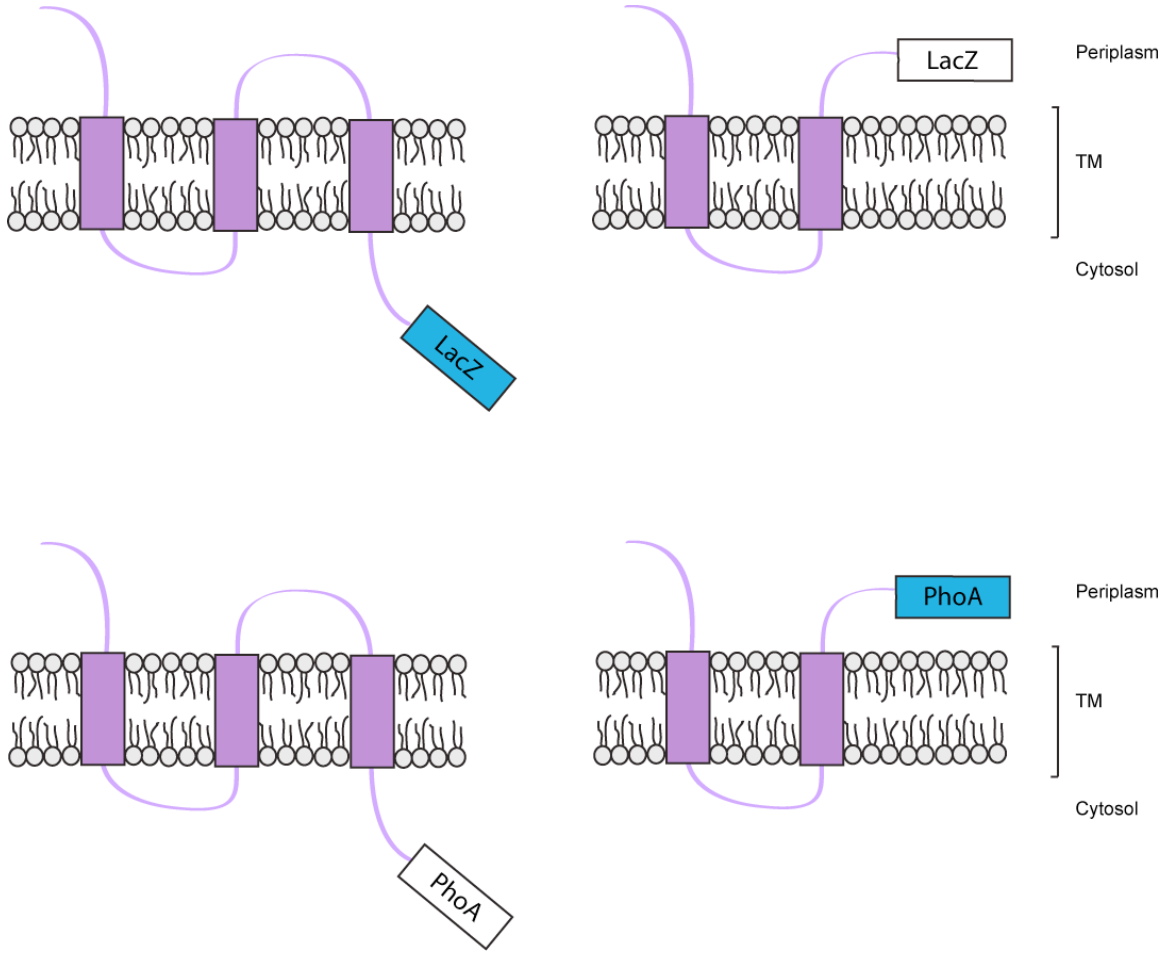


Figure 9. Schematic representation of the enzymatic activity of reporter fusions. Blue indicates when the fusion protein is enzymatically active as the attached hybrid moiety is localized to the correct cellular compartment. Bacteria appear blue when grown on medium containing the appropriate substrate. White indicates an enzymatically inactive fusions and bacteria will appear colourless.

grown on medium containing the substrate analogue X-P. If the PhoA moiety remains in the cytosol then the bacteria appear colourless as PhoA is not enzymatically active. Contrarily, LacZ is only enzymatically active when this moiety of the hybrid protein remains in the cytoplasm. This is visualized when bacteria are grown on medium containing the substrate analog X-Gal and appear blue. Again, if LacZ is translocated to the periplasm (opposite to PhoA) the bacteria will appear colourless, indicating an enzymatically inactive fusion.

The method of reporter fusions can give inconclusive results. This is most likely due to disruption of the tertiary structure of the protein of interest because of truncation or alternate conformations, depending on which reporter protein is present. Correct topological information is only provided if topogenic signals from the target protein are present before the fusion junction to the reporter enzyme [77-80]. Junctions closer to the TM borders and even within TMs can lead to false positive and negative results, making the accurate topological assignment very difficult.

1.6.3 Substituted cysteine accessibility method (SCAM)

The analysis of protein structure has been greatly advanced by a technique which involves the modification of thiol groups in proteins and is designated SCAM. It is a biochemical approach to experimentally determine the topology of inner membrane proteins. SCAM allows the residues on the water-accessible surface of membrane proteins to be systematically mapped. It takes advantage of thiol-reacting molecules, such as maleimides, combined with membrane-permeable and -impermeable reagents that help define the orientation of cytoplasmic residues versus periplasmic residues, in

addition to the length and orientation of TM segments. TM residues are prevented from reacting with maleimides because of their presence in the hydrophobic space so mutating residues around the border of the predicted TM regions can help refine their predicted length [81-87]. In the presence of water, the double bond of the maleimide ring breaks to form an irreversible thiolate (Figure 10). Maleimide reacts 1,000 times more efficiently in a slightly acidic pH range of 6.5-7.5, whereas an increase to above a pH 8 leads to reactivity with primary amines, and hydrolysis of the maleimide ring. Thus, maleimides are particularly suited for this chemistry as they are virtually unreactive until they encounter an available sulfhydryl group [82, 88, 89]. Further, SCAM allows topological assessment under conditions in which inner membrane proteins remain properly folded. A key point of this methodology is the reliance on a cysteineless protein that remains functional, and allows the introduction of single cysteines at specific sites. Single amino acid changes are less likely to disrupt tertiary structure of the protein and alter its native topology, a problem commonly found with reporter fusions. Therefore, SCAM provides the basis for better refinement of the topology of the protein of interest [82, 85, 90, 91].

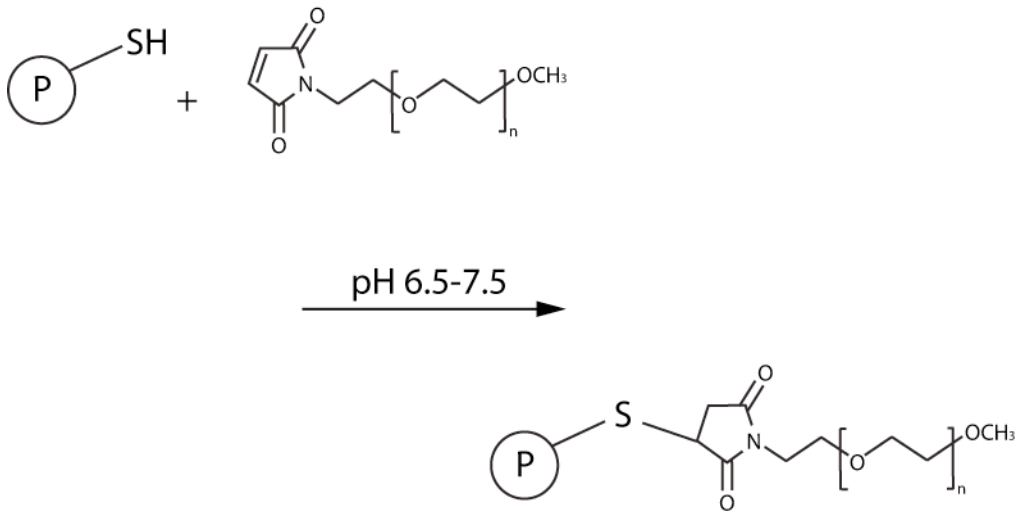


Figure 10. Schematic representation of the sulfhydryl reaction. Maleimide molecules become reactive when they encounter available sulfhydryl groups, such as single cysteines exposed on the protein surface in extra membrane loops. In the presence of water, the double bond of the maleimide ring breaks to form an irreversible thiolate, while the terminal monomethoxy group prevents cross-linking to other available thiols. P = protein of interest; SH = sulfhydryl group (substituted cysteine).

The classical SCAM method employs biotin maleimide and a blocker reagent such as MTSET. Biotin maleimide is internalized by bacteria and labels available cysteines in extramembrane loops, whether in the cytosol or the periplasmic space. MTSET is a positively charged, thiol-reactive reagent that because it cannot cross the inner membrane blocks periplasmic cysteine residues from biotinylation (Figure 11). Cysteine residues in TMs are not labeled due to their hydrophobic environment that prevents access of water molecules that are critical for biotinylation. The difficulty with this strategy is that the biotinylated protein must be purified prior to detection with streptavidin. Purification is critical to separate the modified protein from the background of other native proteins in the membrane that became biotinylated in their native cysteines. Therefore, this classical biotinylation method for SCAM can give considerable background, and is highly labour-intensive and time consuming [20, 51, 92]. An alternative strategy based on PEGylation overcomes this shortcoming and will be discussed in the next section.

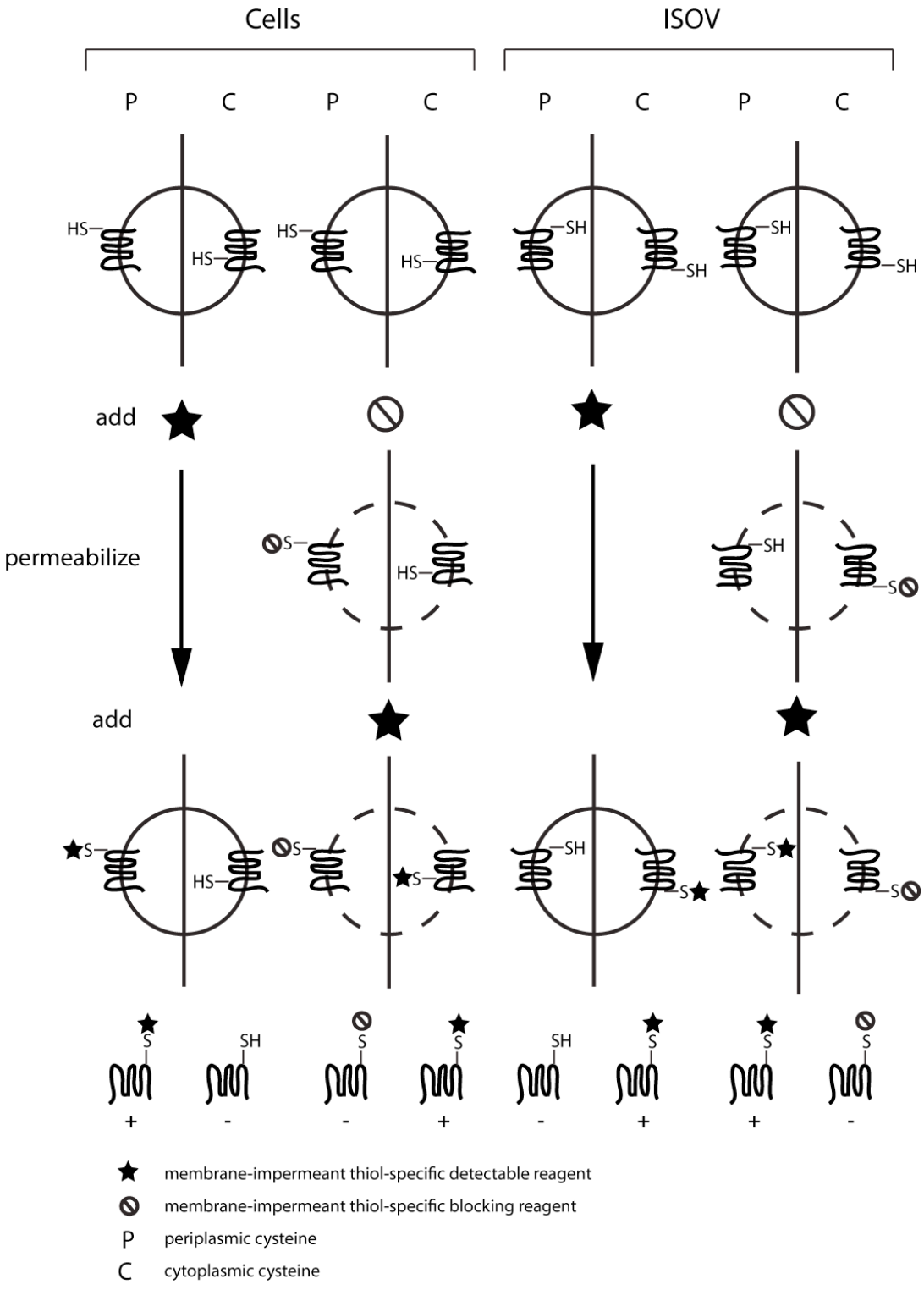

Cells
ISOV
P
C
P
C
P
C
P
C
HS
SH
add
permeabilize
add
S
+
-
-
+
-
+
+
-
membrane-impermeant thiol-specific detectable reagent
membrane-impermeant thiol-specific blocking reagent
periplasmic cysteine
cytoplasmic cysteine

Figure 11. Schematic representation of SCAM (Modified from [82]). A protein of interest containing a single substituted cysteine either to the periplasmic (left half of the circle) or cytoplasmic (right half of the circle) side of the membrane is expressed in cells. Membrane preparations result in inside-out orientated vesicles (ISOV), which are opposite orientations for the same cysteines. Half of the cells or ISOV is reacted with a detectable thiol reagent to label the periplasmic exposed cysteine and the other half is reacted with blocking thiol reagent to protect the periplasmic-exposed cysteines in consequent labeling steps. To specifically label cytoplasmic cysteines, the latter half of cells or ISOV is permeabilized to expose the cytoplasmic cysteines to react with a detectable thiol reagent. The detectable thiol reagent is used to track the labeled target protein from analyzed cells and ISOV. Note ISOV patterns of cysteine labeling (+) and blocking (-) are mirror image of the whole cell patterns.

1.7 PEGylation

PEGylation also employs maleimide reagents, but eliminates the need for protein purification. In this method, the cysteine residues are modified by the covalent attachment of poly(ethylene) glycol (PEG). PEG is a hydrophilic and linear, or branched, polyether terminated with hydroxyl or functional groups. Monomethoxy PEG is the most useful for modification when defining inner membrane protein topologies as it prevents cross-linking from occurring.

PEG-Mal has a molecular mass of 5 kDa and is membrane-impermeable. Using ethylenediaminetetracetic acid (EDTA) to permeabilize the outer membrane, PEG-Mal is able to react with periplasmic residues of the protein in otherwise intact cells. EDTA chelates the divalent metal ions of LPS to create a more permeable outer membrane but still maintains the structure and organization of the bacterial cell envelope [93, 94]. Upon crossing the EDTA-permeabilized outer membrane, PEG-Mal forms an irreversible thiolate with any available periplasmic substituted cysteine. To asses if a substituted cysteine is exposed to the cytoplasm, membrane preparations are performed before PEG-Mal labeling. Membrane preparations give a mix of membrane vesicles where cytoplasmic substituted cysteines become exposed and are therefore available to react with PEG-Mal. It is only TM substituted cysteines, buried in the hydrophobic space, that are unable to form a thiolate with PEG-Mal. The PEG bulk creates an apparent band shift of 10-25 kDa detectable by western blot, indicating that WecA has been modified [89, 93, 95] (Figure 12).

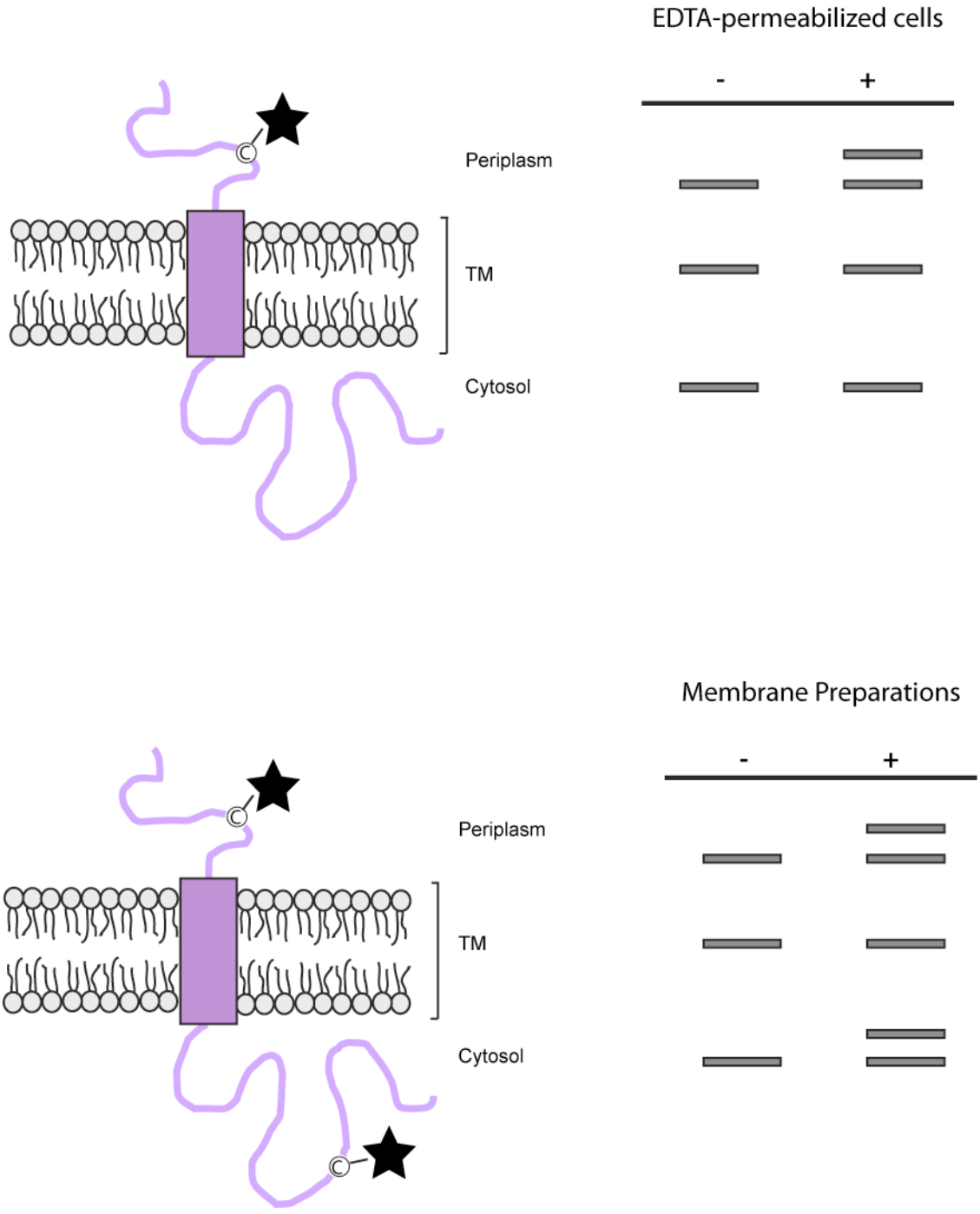

EDTA-permeabilized cells
-
+
Periplasm
TM
Cytosol
Membrane Preparations
-
+
Periplasm
TM
Cytosol
monomethoxy-poly(ethylene) glycol-maleimide (PEG-Mal)

Figure 12. Schematic representation of PEGylation and the theoretical Western blots based on labeling pattern. In EDTA-permeabilized cells only periplasmic-exposed cysteines react with PEG-Mal and show a band-shift. In membrane preparations cysteines on either side of the inner membrane react with PEG-Mal and show a band-shift. TM residues orientated into the lipid bilayer are do not react with PEG-Mal and do not show a band-shift.

1.8 Research objectives

1.8.1 Rationale

The topological model of the *E. coli* WecA was originally derived using the TMHMM computer software and was experimentally refined by SCAM based on biotinylation [20, 82]. From this information WecA was predicted to consist of 11 transmembrane domains (TMs), 5 periplasmic and 5 cytoplasmic loops, with the N-terminus located in the periplasm and the C-terminus in the cytoplasm (Refer to Figure 1). The cytoplasmic location of the C-terminus was independently confirmed by GFP fusion [17, 18, 20].

However, WecA is difficult to over-express and purify, resulting in inconsistencies in the biotin labeling experiments. Furthermore, functional residues detected by mutagenesis are spread over several different cytoloops of the protein [5, 20], making the need to determine the border of the TMs and cytoloops even more critical. Therefore, the central goal of this research project was to determine the boundaries between specific TMs and cytoloops with a reliable thiol labeling method such as PEGylation, which bypasses the need for a purification step.

1.8.2 Specific Objectives

This thesis has two specific objectives:

1) To optimize the PEGylation assay in WecA. To address this objective, I first demonstrated that PEGylation is specific and predictive of the topological location of the targeted residue. The former was accomplished using increasing PEG-Mal concentrations in labeling experiments performed on the cysteineless WecA (negative control) and

WecA mutants with cysteine residues in periplasmic, membrane and cytoplasmic regions, which were previously established by biotin-SCAM. The aim was to demonstrate that PEG-Mal only labels available periplasmic cysteines in intact (EDTA-permeabilized) cells but can also label cytoplasmic exposed cysteines in membrane preparations.

2) To refine the borders between cytoloops 1 and 4, and their corresponding TMs, and to better characterize TMs IV and V. This objective involved creating additional single substituted cysteine mutants in WecA in the areas of interest. Translocation of each mutant derivative to the membrane was confirmed by Western blot and complementation of OAg production in *E. coli* MV501 *in vivo* was verified. Finally, the PEGylation assay was employed to refine the orientation of the substituted residues.

Chapter 2

Material and Methods

2.1 Bacterial strains, growth conditions, and reagents

The bacterial strains and plasmids used in this study are listed in Table 1. Bacteria were grown at 37°C in Luria-Bertani (LB) medium supplemented with ampicillin (100 µg/mL) and/or tetracycline (20 µg/mL) and/or 0.2% (w/v) arabinose, as appropriate. Chemicals were purchased from Sigma (St. Louis, MO, USA), unless otherwise stated. Enzymes were purchased from Agilent Life Sciences and Chemical Analysis (Santa Clara, CA, USA), New England Biolabs (Massachusetts, USA) and Roche Diagnostics (Laval, QC, Canada). Oligonucleotide primers for sequencing and site-directed mutagenesis (listed in Table 2) were purchased from Eurofins MWG Operon (Huntsville, AL, USA). The QIAprep Spin Miniprep Kit by Qiagen (ON, Canada) and the High-Speed Plasmid Mini Kit by Geneaid (New Taipei City, Taiwan) were used to isolate plasmids.

2.2 Site-directed mutagenesis

2.2.1 Plasmid isolation

Plasmids were isolated by a small-scale plasmid DNA preparation. Briefly, 5 mL overnight cultures (16 h at 37° using orbital shaking at 200 rpm) of *E. coli* DH5α containing the appropriate plasmids were pelleted by centrifugation using a 1.5-mL microcentrifuge tube at 16 000 *x g* for one min (repeated until the entire culture were harvested). The supernatant was aspirated and the bacterial cell pellet was resuspended in 200 µL of PD1 Buffer (with RNase A). Two hundred µL of PD2 Buffer was added and the tube was inverted to gently mix the suspension without shearing genomic DNA and incubated at room temperature for at least 2 min, or until the lysate was homogeneous.

Table 1. Bacterial strains and plasmids used in this study

Strain or Plasmid	Relevant Properties	Reference or Source
E. coli Strains		
DH5α	*E. coli* K-12 F^-Φ80*lacZ*Δ M15 *endA recA* *hsdR*(rK^- mK^-) *supE thi gyrA relA* Δ(*lacZYA-argF*)*U169*	Laboratory stock
MV501	VW187; *wecA*::Tn*10* Tc^R	[57]
Plasmids		
pBAD24	Cloning vector	[96]
pJL7	pKV1 expressing cysteineless WecA	[20]
pJL8	pJL7 expressing WecA-D35C	[20]
pJL15	pJL7 expressing WecA-F143C	[20]
pJL19	pJL7 expressing WecA-G181C	[20]
pJL20	pJL7 expressing WecA-R209C	[20]
pJL29	pJL7 expressing WecA-S362C	[20]
pSEF39	pJL7 expressing WecA-Y211C	[51]
pSEF34	pJL7 expressing WecA-V213C	[51]
pSEF35	pJL7 expressing WecA-M215C	[51]
pSEF36	pJL7 expressing WecA-D217C	[51]
pSEF37	pJL7 expressing WecA-S220C	[51]
pSLL1	pJL7 expressing WecA-L4C	This study
pSLL2	pJL7 expressing WecA-S7C	This study
pSLL3	pJL7 expressing WecA-S12C	This study
pSLL4	pJL7 expressing WecA-L19C	This study
pSLL5	pJL7 expressing WecA-A24C	This study
pSLL6	pJL7 expressing WecA-K30C	This study
pSLL7	pJL7 expressing WecA-K41C	This study
pSLL8	pJL7 expressing WecA-L46C	This study
pSLL9	pJL7 expressing WecA-S54C	This study
pSLL10	pJL7 expressing WecA-F61C	This study
pSLL11	pJL7 expressing WecA-V66C	This study
pSLL12	pJL7 expressing WecA-Y69C	This study
pSLL13	pJL7 expressing WecA-A73C	This study
pSLL14	pJL7 expressing WecA-L77C	This study
pSLL15	pJL7 expressing WecA-L83C	This study
pSLL16	pJL7 expressing WecA-R25C	This study
pSLL17	pJL7 expressing WecA-V27C	This study
pSLL18	pJL7 expressing WecA-Q44C	This study
pSLL19	pJL7 expressing WecA-G45C	This study
pSLL20	pJL7 expressing WecA-Q104C	This study
pSLL21	pJL7 expressing WecA-V107C	This study
pSLL22	pJL7 expressing WecA-V110C	This study
pSLL23	pJL7 expressing WecA-M112C	This study

pSLL24	pJL7 expressing WecA-F114C	This study
pSLL25	pJL7 expressing WecA-K116C	This study
pSLL26	pJL7 expressing WecA-Y118C	This study
pSLL27	pJL7 expressing WecA-S120C	This study
pSLL28	pJL7 expressing WecA-L122C	This study
pSLL29	pJL7 expressing WecA-Y124C	This study
pSLL30	pJL7 expressing WecA-L133C	This study
pSLL31	pJL7 expressing WecA-F136C	This study
pSLL32	pJL7 expressing WecA-F139C	This study
pSLL33	pJL7 expressing WecA-T141C	This study
pSLL35	pJL7 expressing WecA-I149C	This study
pSLL38	pJL7 expressing WecA-I47C	This study

Next, 300 µL of PD3 Buffer was added and mixed gently by inversion before centrifuging at 16 000 *x g* for 10 min. The supernatant was added to a PD column, which was placed in a 2-mL collection tube, and was centrifuged at 16 000 *x g* for 30 sec. The flow-through was discarded and 400 µL of W1 Buffer was added to the PD column and then centrifuged at 16 000 *x g* for 30 sec. Again, the flow-through was discarded and 600 µL of Wash Buffer (ethanol added) was added to the PD column and then centrifuged at 16 000 *x g* for 30 sec. The flow-through was discarded one more time before centrifuging the PD column/collection tube at 16 000 *x g* for 3 min to dry the column matrix. The dried PD column was transferred to a clean microcentrifuge tube and 50 µL of Elution Buffer was added to the centre of the column matrix. After incubation at room temperature for at least 2 min, or until the Elution Buffer was completely absorbed, the tube was centrifuged at 16 000 *x g* for 5 min to elute the DNA.

2.2.2 Agarose gel electrophoresis

DNA fragments were separated by agarose gel electrophoresis on 0.8% agarose gel (w/v) prepared in 1X TBE buffer (90 mM Tris-Borate, 1 mM EDTA [pH 7.5]) supplemented to a final concentration of 0.0005 mg/mL ethidium bromide. DNA samples were loaded with 10X DNA agarose loading dye (50% glycerol [v/v], 0.1% bromophenol blue [w/v], 100 mM EDTA [pH 7.5]), and the gels were subjected to electrophoresis at 100 V. A 1-kilobase pair (kb) or 100-base pair (bp) molecular weight GeneRule DNA Ladder Marker (MBI Fermentas, Burlington, ON, Canada) was loaded in each gel as DNA fragment size reference. DNA gels were visualized using a GelDoc 2000 Imaging Unit and digitally processed with Multi-Analyst software (Bio-Rad, Hercules, CA, USA).

2.2.3 Polymerase Chain Reaction (PCR) with Pfu Turbo DNA polymerase

WecA derivatives with single amino acid replacements were constructed with the Quik-Change Site Directed Mutagenesis Kit by Stratagene (Agilent Life Sciences and Chemical Analysis [Santa Clara, CA, USA]). The forward and reverse primer pairs (Table 2) were designed to create the plasmids pSLL1 through pSLL38 encoding the proteins listed in Table 1. The 38 individual mutations were made in the parental $WecA_{FLAG\text{-}7xHis}$ gene (plasmid pJL7) expressing cysteineless WecA. See figure 13 for the locations of the mutations indicated on the topological map of WecA. The PCR reactions were prepared by mixing 1 μL of pJL7 template, 1 μL of dNTP mix (10 mM), 5 μL of 10X Pfu Turbo reaction buffer, 39.5 μL of double-distilled water (ddH_2O), and 1.5 μL of the forward and reverse primers of choice. Then, 0.5 μL of Pfu Turbo DNA polymerase was added to each reaction volume (~50 μL) and PCR thermal cycling was performed using a Bio-Rad Dyad Peltier Thermal Cycler. The PCR program involved an initial 30 sec denaturation at 95°C followed by 16 subsequent cycles of denaturation at 95°C for 30 sec, primer annealing at 55°C for 1 min, and extension at 68°C for 12 min. The PCR reactions were digested overnight with 1 μL of the restriction enzyme DpnI at 37°C to degrade any residual pJL7 plasmid DNA. Successful PCR amplification of pre- and post-DpnI PCR products were verified by agarose gel electrophoresis as described above in section 2.2.2. Next, 5 μL of confirmed products were introduced by transformation into chemically competent DH5α *E. coli* as described below in section 2.4.1. Plasmids were isolated from transformed colonies and the correct mutation confirmed by DNA sequencing using the oligonucleotide primers 252 and 258 (Table 2). Colonies with the

Table 2. Oligonucleotide primers used in this study

WecA strain and mutants	Primer	Sequence*
$WecA_{L4C}$	6325	5’ – ACATATGAATTTA**TGT**ACAGTGAGTACT
	6326	5’ – AGTACTCACTGT**ACA**TAAATTCATATGT
$WecA_{S7C}$	6327	5’ – TTACTGACAGTG**TGT**ACTGATCTCATC
	6328	5’ – GATGAGATCAGT**ACA**CACTGTCAGTAA
$WecA_{S12C}$	6329	5’ – ACTGATCTCATC**TGT**ATTTTTTTATTC
	6330	5’ – GAATAAAAAAAT**ACA**GATGAGATCAGT
$WecA_{L19C}$	6331	5’ – TTATTCACGACA**TGT**TTTCTGTTTTTT
	6332	5’ – AAAAAACAGAAA**ACA**TGTCGTGAATAA
$WecA_{A24C}$	6333	5’ – TTTCTGTTTTTT**TGT**CGTAAGGTGGCA
	6334	5’ – TGCCACCTTACG**ACA**AAAAAACAGAAA
$WecA_{R25C}$	6650	5’ – TCTGTTTTTTGCC**TGT**AAGGTGGCAAAA
	6651	5’ – TTTTGCCACCTT**ACA**GGCAAAAAACAGA
$WecA_{K26C}$	6861	5’ – TTTTTTGCCCGT**TGT**GTGGCAAAAAAA
	6862	5’ – TTTTTTTGCCAC**ACA**ACGGGCAAAAAA
$WecA_{V27C}$	6652	5’ – TTTGCCCGTAAG**TGT**GCAAAAAAAGTC
	6653	5’ – GACTTTTTTTGC**ACA**CTTACGGGCAAA
$WecA_{K30C}$	6335	5’ – AAGGTGGCAAAA**TGT**GTCGGTTTAGTG
	6336	5’ – CACTAAACCGAC**ACA**TTTTGCCACCTT
$WecA_{K41C}$	6337	5’ – CCAAACTTCCGC**TGT**CGTCACCAGGGA
	6338	5’ – TCCCTGGTGACG**ACA**GCGGAAGTTTGG
$WecA_{Q44C}$	6654	5’ – CGCAAACGTCAC**TGT**GGATTGATACCT
	6655	5’ – AAGTATCAATCC**ACA**GTGACGTTTGCG
$WecA_{G45C}$	6656	5’ – AAACGTCACCAG**TGT**TTGATACCTCTC
	6657	5’ – GAGAGGTGTCAA**ACA**CTGGTGACGTTT
$WecA_{L46C}$	6359	5’ – CGTCACCAGGGA**TGT**ATACCTCTCGTT
	6360	5’ – AACGAGAGGTAT**ACA**TCCCTGGTGACG
$WecA_{I47C}$	6863	5’ – CACCAGGGCTTG**TGT**CCTCTCGTTGGG
	6864	5’ – CCCAACGAGAGG**ACA**CAATCCCTGGTG

$WecA_{S54C}$	6361	5’ – GTTGGGGGGATT**TGT**GTTTACGCAGGG
	6362	5’ – CCCTGCGTAAAC**ACA**AATCCCCCCAAC
$WecA_{F61C}$	6363	5’ – GCAGGGATTGCC**TGT**ACGTTCGGAATT
	6364	5’ – AATTCCGAACGT**ACA**GGCAATCCCTGC
$WecA_{V66C}$	6365	5’ – ACGTTCGGAATT**TGT**GATTACTATATT
	6366	5’ – AATATAGTAATC**ACA**AATTCCGAACGT
$WecA_{Y69C}$	6367	5’ – ATTGTCGATTAC**TGT**ATTCCGCATGCA
	6368	5’ – TGCATGCGGAAT**ACA**GTAATCGACAAT
$WecA_{A73C}$	6369	5’ – TATATTCCGCAT**TGT**TCTCTCTATCTC
	6370	5’ – GAGATAGAGAGA**ACA**ATGCGGAATATA
$WecA_{L77C}$	6371	5’ – GCATCTCTCTAT**TGT**GCTGCTGCCGGT
	6372	5’ – ACCGGCAGCAGC**ACA**ATAGAGAGATGC
$WecA_{L83C}$	6373	5’ – GCTGCCGGTGTG**TGT**GTTTTCATTGGC
	6374	5’ – GCCAATGAAAAC**ACA**CACACCGGCAGC
$WecA_{Q104C}$	6658	5’ – CGTGCCACCATA**TGT**GCCGCTGTTGGC
	6659	5’ – GCCAACAGCGGC**ACA**TATGGTGGCACG
$WecA_{V107C}$	6660	5’ – ATACAGGCCGCT**TGT**GGCATTGTTATG
	6661	5’ – CATAACAATGCC**ACA**AGCGGCCTGTAT
$WecA_{V110C}$	6662	5’ – GCTGTTGGCATT**TGT**ATGATGGTGTTC
	6663	5’ – GAACACCATCAT**ACA**AATGCCAACAGC
$WecA_{M112C}$	6699	5’ – GGCATTGTTATG**TGT**GTGTTCGGCAAG
	7000	5’ – CTTGCCGAACAC**ACA**CATAACAATGCC
$WecA_{F114C}$	6664	5’ – GTTATGATGGTG**TGT**GGCAAGCTTTAT
	6665	5’ – ATAAAGCTTGCC**ACA**CACCATCATAAC
$WecA_{K116C}$	6701	5’ – ATGGTGTTCGGC**TGT**CTTTATCTCAGT
	6702	5’ – ACTGAGATAAAG**ACA**GCCGAACACCAT
$WecA_{Y118C}$	6666	5’ – TTCGGCAAGCTT**TGT**CTCAGTAGCCTG
	6667	5’ – CAGGCTACTGAG**ACA**AAGCTTGCCGAA
$WecA_{S120C}$	6703	5’ – AAGCTTTATCTC**TGT**AGCCTGGGTTAT
	6704	5’ – ATAACCCAGGCT**ACA**GAGATAAAGCTT

$WecA_{L122C}$	6705	5’ – TATCTCAGTAGC**TGT**GGTTATATCTTT
	6706	5’ – AAAGATATAACC**ACA**GCTACTGAGATA
$WecA_{Y124C}$	6707	5’ – AGTAGCCTGGGT**TGT**ATCTTTGGCTCC
	6708	5’ – GGAGCCAAAGAT**ACA**ACCCAGGCTACT
$WecA_{G127C}$	6865	5’ – GGTTATATCTTT**TGT**TCCTGGGAGATG
	6866	5’ – CATCTCCCAGGA**ACA**AAAGATATAACC
$WecA_{L133C}$	6709	5’ – TGGGAGATGGTG**TGT**GGACCGTTTGGT
	6710	5’ – ACCAAACGGTCC**ACA**CACCATCTCCCA
$WecA_{F136C}$	6668	5’ – GTGCTCGGACCG**TGT**GGTTACTTCCTG
	6669	5’ – CAGGAAGTAACC**ACA**CGGTCCGAGCAC
$WecA_{F139C}$	6670	5’ – CCGTTTGGTTAC**TGT**CTGACGCTATTT
	6671	5’ – AAATAGCGTCAG**ACA**GTAACCAAACGG
$WecA_{T141C}$	6711	5’ – GGTTACTTCCTG**TGT**CTATTTGCCGTC
	6712	5’ – GACGGCAAATAG**ACA**CAGGAAGTAACC
$WecA_{I149C}$	6674	5’ – GTCTGGGCGGCC**TGT**AATGCGTTCAAC
	6675	5’ – GTTGAACGCATT**ACA**GGCCGCCCAGAC

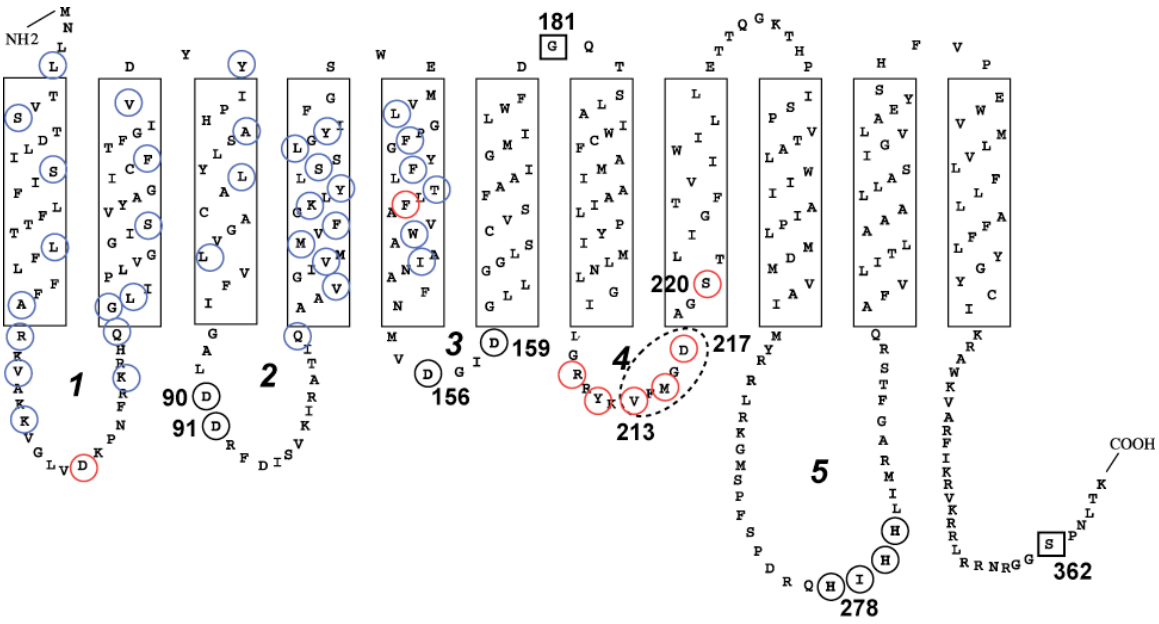


Figure 13. Additional amino acid substitution locations in WecA (Modified from [51]). The WecA topological model originally predicted with TMHMM and experimentally refined using SCAM was used as a guideline for the selection of amino acids to mutate to cysteine. Black circles indicate critical residues previously determined [22, 23, 69]. Black squares indicate the periplasmic (G181C) and cytoplasmic (S362C) controls. Blue circles indicate mutations created in this study. Red circles indicate previous mutations created that are tested in this study [20, 51, 92]. Bolded numbers indicate cytoplasmic loops.

plasmids containing the desired amino acid replacements were stored in stock solution (20% glycerol [v/v], 20% LB [w/v], ddH_2O) at -80°C.

2.2.4 DNA sequencing

DNA sequencing was performed by the Core Molecular Biology Facility, York University, ON, Canada and/or Eurofins MWG Operon, Huntsville, AL, USA.

2.3 Competent cell preparation

2.3.1 DH5α chemically competent cells

E. coli DH5α chemically competent cells were made by the rubidium chloride ($RbCl_2$) method adapted from Sambrook *et al*. [97]. From a frozen stock, a 5 mL overnight culture of DH5α cells was grown. The overnight culture was added to 150 mL of Super Optimal Broth (SOB) supplemented with 1.5 mL of 2 M Mg^{2+} and incubated until an optical density (OD_{600}) of 0.6 to 0.8 was reached. The culture was transferred to sterile 25-mL plastic Nalgene tubes and chilled on ice for 15 min. Cells were harvested by centrifugation at 5927 *x g* for 10 min at 4°C. The pellet was resuspended in 50 mL of Buffer 1 (100 mM $RbCl_2$, 50 mM $MnCl_2 \cdot 4H_2O$, 30 mM potassium acetate [pH 7.5], 10 mM $CaCl_2 \cdot 2H_2O$, and 15% glycerol [v/v] adjusted to a final pH of 5.8 with 0.2 M acetic acid), placed on ice for 15 min, and centrifuged again. The pellet was resuspended in 12 mL of Buffer 2 (10 mM $RbCl_2$, 75 mM $CaCl_2 \cdot 2H_2O$, 10 mM MOPS, and 15% glycerol [v/v] adjusted to a final pH of 6.8 with NaOH) and placed on ice for another 15 min. Finally, cells were aliquoted into 200 μL fractions in 1.5 mL microcentrifuge tubes and

were flash-frozen in an ethanol/dry ice bath. Chemically competent DH5α cell were stored at -80°C until needed.

2.3.2 MV501 electrocompetent cells

Starting from a frozen stock, a 5 mL overnight culture of MV501 cells was grown supplemented with tetracycline. The overnight was inoculated into 25 mL of LB with an adjusted optical density (OD_{600}) of 0.1 and grown for approximately 1 h, until the optical density (OD_{600}) reached 0.5 to 0.6. The culture was then chilled on ice for 10 min and transferred to a sterile 25-mL plastic Nalgene tube. The cells were harvested by centrifugation at 5927 *x g* for 10 min at 4°C. The supernatant was decanted and the pellet was resuspended in 5 mL of cold, sterile ddH_2O and centrifuged again at 5927 *x g* for 10 min at 4°C. Next, the supernatant was decanted and the pellet was resuspended in 1 mL of cold, sterile ddH_2O before being divided into two aliquots in 1.5-mL microcentrifuge tubes, and topped up to 1 mL with cold, sterile ddH_2O. The cells were harvested in the tabletop centrifuge at 16 000 *x g* and resuspended in 1 mL of cold, sterile ddH_2O. This step was repeated once more before the final pellet was resuspended in 500 µL of 20% glycerol/ddH_2O. Electrocompetent MV501 cells were aliquoted into 50 µL volumes and frozen at -80°C until needed.

2.4 Transformation of *E. coli* strains

2.4.1 Transformation of DH5α chemically competent cells

Chemically competent cells were thawed on ice and 5 µL of the PCR product was added and mixed gently. The cells were left on ice for 30 min before heat-shock at 42°C

for 2 min in a water bath. Then, 800 µL of SOC media was added to each aliquot and the cells were incubated at 37°C for 2 h. One hundred twenty five µL of the recovered cell suspension was spread on LB-Agar plates supplemented with ampicillin and left overnight at 37°C.

2.4.2 Transformation of MV501 electrocompetent cells

Electrocompetent cells were thawed on ice and 2 µL of the appropriate plasmid was added and mixed gently. The aliquots were transferred to chilled electroporation cuvettes (Bio-Rad, CA, USA) and shocked at 1.25V using a Bio-Rad Gene Pulser to obtain a time constant of at least 4. The cells were recovered in 800 µL of SOC medium at 37°C for 2 h in a water bath. Then, 100 µL of the recovered cell suspension was spread on LB-Agar plates supplemented with ampicillin and tetracycline and left overnight at 37°C.

2.5 Protein expression

2.5.1 Isolation of total membrane proteins

Five-mL cultures of MV501 transformed with a plasmid containing parental or mutated *wecA* were incubated overnight at 37°C in LB supplemented with ampicillin and tetracycline. The overnight cultures were inoculated into 25 mL flasks of LB with an adjusted optical density (OD_{600}) of 0.1 and grown for approximately 1.5 h, until mid-logarithmic phase, where the optical density (OD_{600}) reached 0.5 to 0.6. At this time point, transcription of the plasmid-encoded *wecA* was induced by arabinose at a final concentration of 0.2%, followed by incubation at 37°C for 3 h. Cells were harvested by

centrifugation at 9 300 *x g* for 10 min at 4°C, washed with 0.9% NaCl, and pelleted again at 9 300 *x g* for 10 min at 4°C. The cells were resuspended in 2 mL of 50 mM TAE buffer (50 mM Tris-acetate [pH 8.5], 1 mM EDTA), supplemented with 80µL of Complete-EDTA-free protease inhibitor, transferred to 15-mL Falcon tubes, and lysed with the Constant Systems Ltd (UK) cell disruptor set to the "one shot" mode at 27 000 Psi. Cell debris was pelleted by centrifugation at 17 418 *x g* for 15 min at 4°C. Next, the supernatant was transferred to microcentrifuge tubes and total membranes were pelleted by ultracentrifugation at 30 966 *x g* for 35 min at 4°C. The membranes were resuspended in 10 µL of TAE buffer and total protein concentration of the fractions was measured immediately (see below).

2.5.2 Quantification of total membrane proteins

The concentration of protein in the total membrane fractions was measured using the Bio-Rad Protein Assay, a modified Bradford Assay. One µL of membrane preparation was diluted into 799 µL of ddH_2O and 200 µL of Bio-Rad Protein Assay Dye for a final volume of 1 mL. Similarly, standards containing 0, 2, 4, 6, 8, and 10 µg of bovine serum albumin (BSA) were prepared. The absorbance of the standards and the membrane preparations was measured at 595 nm. The values obtained from the standards were used to construct a standard curve from which total protein concentration of each membrane preparation was extrapolated.

2.5.3 Detection of WecA and WecA mutant derivatives

Wild type WecA and the WecA mutants contained a C-terminal FLAG epitope so that the proteins could be detected by Western blot analysis. A previously described protocol was adapted to first resolve the proteins by gel electrophoresis [98]. Volumes of membrane preparations containing 5-10 μg of total protein were combined with 3X protein dye (0.875 M Tris [pH 6.8], 6% [w/v] SDS, 30% [v/v] glycerol, 0.03% [w/v] bromophenol blue, 15% [v/v] β-mercaptoethanol) and separated by 14% Tris-glycine SDS-PAGE (Running Gel: 14% ABA 44:0.8, 0.2% SDS, 0.375 M Tris-HCl [pH 8.0], 0.025% ammonium persulfate [APS], 0.025% N,N,N',N'-tetramethylehtylenediamine [TEMED]; Stacking Gel: 4.5% ABA 44:0.8, 0.1% SDS, 0.125 M Tris-HCl [pH 6.8], 0.025% APS, 0.025% TEMED) electrophoresis in 1X SDS buffer (25 mM Tris [pH 7.6], 200 mM glycine, 2.3 mM SDS) using the Mini-Protean 3 Cell (Bio-Rad, CA, USA) at 110-120V until the dye front ran through the stacking gel then increased to 130-150V. A pre-stained Broad-Range Molecular Weight Marker (Bio-Rad) was included on all gels as a molecular mass standard. Electrophoresis was performed until the dye front reached the bottom of the protein gel.

Next, using the Bio-Rad Transblot system, proteins were transferred to nitrocellulose membranes (Pall Life Sciences, Pensacola, FL). Briefly, the protein gel, a nitrocellulose membrane, & blotting papers (VWR, UK), both cut to the size of the gel were assembled as described elsewhere [97]. The transfer was performed in chilled Transfer Tris-Glycine Buffer (25 mM Tris, 190 mM glycine, 0.1% SDS, 20% methanol) at 250 mA for 1 h. After disassembling the transfer system, the nitrocellulose membranes were washed with Tris Buffered Saline (TBS) (1 M Tris-HCl [pH 7.7] and 5 M NaCl).

Then, the membranes were blocked overnight in Western Blocking Reagent (Roche) diluted in TBS in a cold room. The following day, the membranes were rinsed with TBS then incubated for at least 1.5 h in TBS with 1:10 000 Anti-FLAG M2 monoclonal (mouse) antibody (Sigma, MO, USA) by rocking at room temperature. Next, the membranes were washed with TBS three times for 10 min each by rocking at room temperature. The membranes were incubated for at least 15 min in TBS with 1:20 000 Alexa Fluor 680 anti-mouse antibody and 700µL of Western Blocking Reagent by rocking at room temperature. After 3 more washed with TBS, the membranes were developed using the Odyssey V3.0 infrared imagining by Licor (Li-Cor, Lincoln, NB).

2.6 LPS production

2.6.1 LPS isolation

Previously described protocols were modified to isolate LPS from whole bacterial cells [99, 100]. Parental MV501 or derivatives transformed with a plasmid parental or mutated *wecA* were grown overnight as 5 mL cultures in LB supplemented with ampicillin and tetracycline and 0.002% arabinose. The optical density (OD_{600}) was measured and the turbidity was adjusted to 2 in a final volume of 1.5 mL in 1X PBS. From the normalized cell suspensions, whole cells were harvested by centrifugation at 16 000 *x g* in a tabletop centrifuge for 1 min at room temperature. The pellets were resuspended in 150 µL of LPS Lysis Buffer (2% SDS, 4% β-mercaptoethanol, 10% glycerol, 1 M Tris [pH 6.8], 0.01% bromophenol blue) and boiled for 10 min. Ten µL of proteinase K (20 mg/mL) was added to the lysate, followed by vortexing and incubation at 60°C in a water bath for 2 h. Then, 150 µL of pre-warmed 95% phenol was added and

each sample was incubated at 70°C in a water bath for 15 min with vortexing every 5 min. Samples were then cooled on ice for 15 min before centrifugation at 16 000 *x g* for 10 min at room temperature to separate the aqueous phase containing the free LPS from the phenol phase containing any residual cellular components. Next, the clear phase was transferred to a clean microcentrifuge tube and 500 µL of ether was added to remove any residual phenol by mixing it with the aqueous phase by gentle inversion. The samples were centrifuged at 16 000 *x* g for 1 min at room temperature and the ether phase was aspirated. LPS preparations were stored at -20°C.

2.6.2 Gel electrophoresis (Tricine SDS-PAGE)

Three µL of LPS was mixed with 3 µL of 3X protein dye and resolved by gel electrophoresis in a 14% Tricine SDS-PAGE (Running Gel: 28.4% A-6-BA [49.5% acrylamide-6% bisacrylamide], 33.2% gel buffer (3 M Tris [pH 8.5], 0.3% SDS), 10.4% [v/v] glycerol, 0.2% APS, and 0.12% TEMED; Stacking Gel: 3.9% A-3BA [49.5% acrylamide-3% bisacrylamide], 24.5% gel buffer, 1% APS, and 0.44% TEMED) using the Mini-Protean 3 Cell (Bio-Rad). Gels were run at 50V through the stacking gel (~ 30 min) and 130V through the running gel until it ran off (~ 2 h), then run for 20 more min in a two-buffer system (anode buffer (0.2 M Tris [pH 8.9]) and cathode buffer (0.1 M Tris [pH 8.25], 0.1% SDS)).

2.6.3 LPS Silver Staining

After electrophoresis, gels were fixed overnight in a 60% methanol, 10% acetic acid solution in order to visualize the LPS using a modified version of a previously

described silver stain protocol [101]. Gels were rehydrated in a 7.5% acetic acid solution with rocking for 30 min at room temperature. Next, the gels were oxidized in a 0.2% solution of periodic acid with rocking for 30 min at room temperature. The gels were washed three times with ddH_2O for 10 min. A silver staining solution (0.18% sodium hydroxide, 1.4% ammonium hydroxide and 19.4% silver nitrate) was applied for 15 min followed by three more washed with ddH_2O for 10 min. The gels were developed in a 0.05% citric acid, 10% methanol, 0.019% formaldehyde solution until the LPS banding pattern appeared. The developing process was stopped quickly by three more washes with ddH_2O for 10 min. Images of LPS banding patterns were created by scanning the gels using the HP Scanjet 4070 Photosmart Scanner.

2.7 PEGylation

2.7.1 Preparation of EDTA-permeabilized cells for labeling

Five-mL cultures of MV501 transformed with a plasmid containing parental or mutated *wecA* were incubated overnight at 37°C in LB supplemented with ampicillin and tetracycline. The overnight cultures were inoculated into 100 mL flasks of LB supplemented again with ampicillin and tetracycline with an adjusted optical density (OD_{600}) of 0.1 and grown for approximately 1.5 h, until mid-logarithmic phase, where the optical density (OD_{600}) reached 0.5 to 0.6. At this time point, transcription (protein expression) of the plasmid-encoded *wecA* was induced by arabinose at final concentration of 0.2% and the cells were further incubated at 37°C for 3 h. Whole cells were harvested by centrifugation at 9 300 *x g* for 10 min at 4°C, washed with 5 mL of HEPES/$MgCl_2$ Buffer (50 mM HEPES [pH 6.8], 5 mM $MgCl_2$), and pelleted again at 9 300 *x g* for 10

min at 4°C. At this point, the pellets were resuspended in 1 mL of HEPES/$MgCl_2$ and the optical density (OD_{600}) was measured, before being aliquoted in microcentrifuge tubes into volumes with a normalized OD of 5. The cells were pelleted again in at 16 000 *x g* for 1 min then resuspended in 100µL of HEPES/$MgCl_2$ Buffer with 5 mM of 0.5 M EDTA and 1 mM of 10 mM PEG-Mal. The reaction was allowed to proceed for one h at room temperature before being quenched with 45 mM of 0.5 mM DTT for 10 min at room temperature. The labeled cells were pelleted by centrifugation for 5 min, the supernatant was removed, and the pellets were frozen at -20°C until use. Cell pellets were thawed on ice before being resuspended in 1 mL of HEPES/$MgCl_2$ Buffer supplemented with 40 µL of Complete EDTA-free protease inhibitor (Roche). Cells were lysed and total membrane preparations were collected following the protocol described above (section 2.5.1) and resuspended in 10 µL of HEPES/$MgCl_2$ Buffer.

2.7.2 Preparation of total membrane fractions for labeling

Five-mL cultures of MV501 transformed with a plasmid containing parental or mutated *wecA* were incubated overnight at 37°C in LB supplemented with ampicillin and tetracycline. The overnight cultures were inoculated into 100 mL flasks of LB supplemented again with ampicillin and tetracycline with an adjusted optical density (OD_{600}) of 0.1 and grown for approximately 1.5 h, until mid-logarithmic phase, where the optical density (OD_{600}) reached 0.5 to 0.6. At this time point, transcription (protein expression) of the plasmid-encoded *wecA* was induced by arabinose at final concentration of 0.2% and the cells were further incubated at 37°C for 3 h. Whole cells were harvested by centrifugation at 9 300 *x g* for 10 min at 4°C, washed with 5 mL of HEPES/$MgCl_2$

Buffer and pelleted again at 9 300 *x g* for 10 min at 4°C. At this point, the pellets were resuspended in 1 mL of HEPES/$MgCl_2$ and the optical density (OD_{600}) was measured, before being aliquoted in microcentrifuge tubes into volumes with a normalized OD of 5. The cells were pelleted again in at 16 000 *x g* for 1 min then the supernatant was removed and the pellets were frozen at -20°C until use. Cell pellets were thawed on ice before being resuspended in 1 mL of HEPES/$MgCl_2$ Buffer supplemented with 40 µL of Complete EDTA-free protease inhibitor (Roche). Cells were lysed and total membrane preparations were collected following the protocol described above (section 2.5.1). The membranes were resuspended in 20 µL aliquots of HEPES/$MgCl_2$ Buffer with 1 mM of 10 mM PEG-Mal and the reaction was allowed to proceed at room temperature for one h before being quenched with 45 mM of 0.5 M DTT for 10 min. The labeled membranes were pelleted by centrifugation for 5 min and resuspended in 10 µL of HEPES/$MgCl_2$ Buffer.

2.7.3 Detection of PEGylated WecA

Parental WecA and the WecA mutants were detected by Western blot as described above (section 2.5.3). PEGylated WecA resulted in a band shift of approximately 10- to 25-kDa gain in apparent mass relative to parental WecA, producing a characteristic double-band (see figure 11 again for reference).

2.8 Statistics

The data and figures are representative of typical experiments which were repeated at least twice in each case.

Chapter 3

Results

3.1 Rationale for this study

The current topological model of the *E. coli* WecA encompasses 11 TMs, five periplasmic and five cytoplasmic loops, with the N-terminus located in the periplasm and the C-terminus in the cytoplasm [17, 18, 20]. Several residues in cytoloops 2, 3 and 4 are conserved among PNPT family members. Cytoloop 5, the least conserved, is thought to recognize the carbohydrate moiety of the nucleotide sugar substrate (Figure 7). This model was derived from a combination of *in silico* approaches and biotin-SCAM [20, 51, 92]. The classical biotin-SCAM method relies on protein purification to separate the protein of interest from other native, unmodified proteins that contain exposed cysteines, which can also be labeled. Because the WecA protein is difficult to purify we employed SCAM combined with PEGylation. In this method, the cysteine residues are modified by the covalent attachment of poly(ethylene) glycol and the labeled proteins can be readily detected by Western blot by an increase in their apparent molecular mass. This approach simplifies the analysis of cysteine-replaced mutant WecA proteins.

This research was aimed at optimizing the PEGylation protocol for WecA and defining the borders between cytoloops 1 and 4, and their corresponding TMs, as these cytoloops contain residues that are important for enzymatic activity. Moreover, TMs IV and V also have residues predicted to interact with Und-P.

3.2 Expression of WecA mutant enzymes in total membrane fractions

Amino acid substitutions, especially in the TMs, can lead to improper protein folding and degradation. Therefore, SDS-PAGE and Western Blot were used to verify that the mutant proteins constructed in this work are properly targeted to the membrane and expressed at similar levels as the parental WecA. Typically, mutant proteins that cannot properly insert in the membrane cannot be detected in membrane preparations or they are barely detectable, suggesting protein degradation. Analysis of total membrane fractions from bacteria expressing the 38 single-cysteine WecA mutants constructed here were detectable by Western blot, revealing a polypeptide with an apparent mass of 36 kDa (Figure 14). These experiments also revealed a higher molecular weight band corresponding to a WecA aggregate that is routinely seen under our experimental conditions due to the mild denaturation required for visualization of membrane proteins [69]. The single-cysteine mutant proteins were expressed at comparable levels to that of parental WecA, although some variation between mutants can be attributed to loading differences.

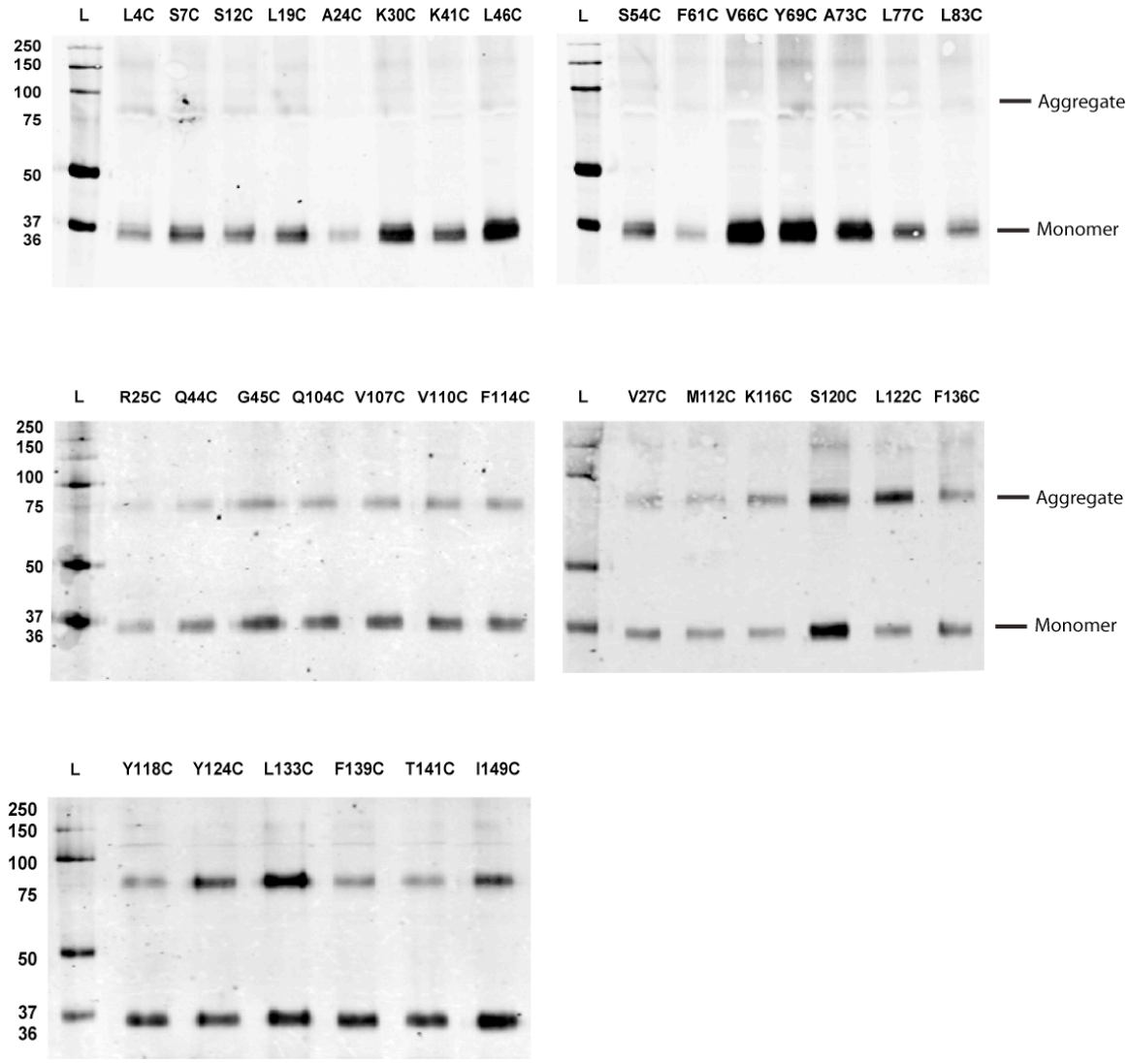


Figure 14. Expression of WecA mutant derivatives. Total membrane fractions were prepared from *E. coli* MV501 transformed with plasmids encoding *wecA* mutants. Bacteria were grown until mid-exponential phase and protein expression was induced by arabinose. Cells were harvested, lysed, and the membrane protein fractions were isolated by ultracentrifugation, and a sample of each mutant (5-10μg of protein) was separated by 14% SDS-PAGE. Proteins were transferred to a nitrocellulose membrane, and WecA was detected by anti-FLAG, with Alexa Fluor 680 anti-mouse used as secondary antibody. The positions of the molecular mass markers are indicated on the left, the mutations on top (L=ladder), and the protein monomers and aggregates on the right.

3.3 Complementation of O-antigen production by WecA mutant enzymes

A protein that is topologically correct should also be functional unless the Cys either replaces a catalytic residue or causes a localized structural perturbation affecting enzyme function. Therefore, we examined the effects of the single-cysteine mutations on OAg production by *in vivo* complementation of *E. coli* strain MV501 containing a *wecA*::Tn*10* insertion. MV501 competent cells were transformed with plasmids encoding *wecA* or one of the *wecA* substituted cysteine mutants. LPS was isolated from the transformants and visualized by silver staining as described in Materials and Methods (Figure 15). The amount of OAg produced by each mutant was compared to parental WecA. K41C, A73C, Q104C, L122C, F136C, and F139C WecA proteins mediated much less polymeric OAg, while M112C, K116C, and T141C fail to complement. These differences, compared to parental WecA (pJL7), would suggest diminished *in vivo* function. However, the differences could also be due to loading differences, as the mutant proteins are stably inserted in the membrane. Similar observations were made with the previously constructed mutants D35C and F143C [20, 92].

Furthermore, similar results were also observed with the replacement of various aromatic residues with conservative, polar mutations located in TMs IV and V, the putative polyisoprenol recognition sequence (PIRS motif). The PIRS motif is predicted to be in transmembrane spanning regions and involved in specific interactions between the hydrophobic residues of the motif and the isoprenoid [5, 102]. If Und-P interacts with WecA at this site by multiple hydrophobic interactions, then increasing the polarity of the contact region should result in a proportional decrease in polymeric OAg production. While the mutants constructed by E. Haggerty, a previous M.Sc. student in our

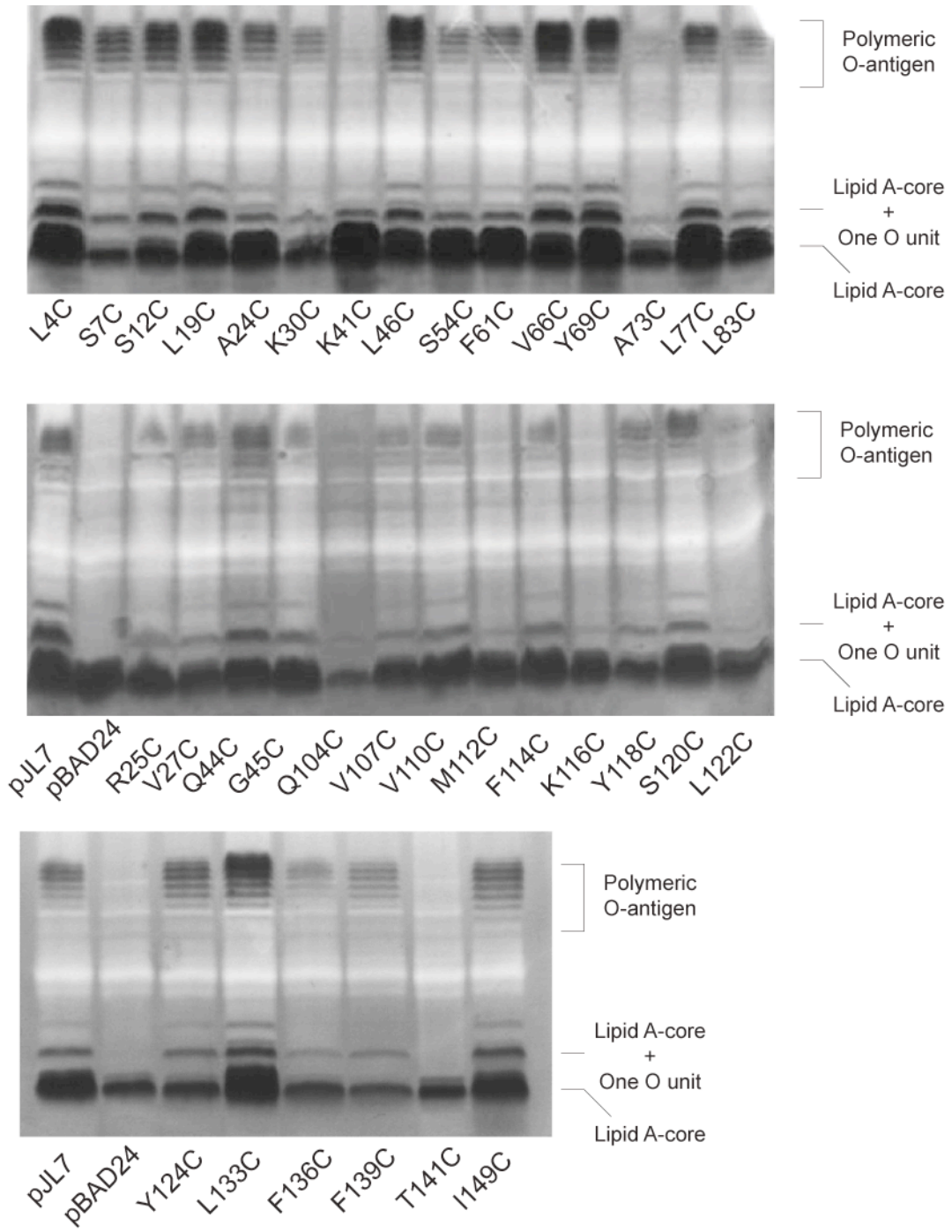


Figure 15. Complementation of OAg production by wild-type WecA and mutant derivatives. LPS was isolated from *E. coli* MV501 transformed with a plasmid encoding either *wecA* or a *wecA* mutant. LPS structural components were resolved by electrophoresis on a 14% Tricine SDS-PAGE, which was fixed overnight and silver stained. The plasmids pJL7 and pBAD24 represent the positive and negative controls, respectively.

laboratory, were not cysteine-substitutions, the polar conservative mutations, such as phenylalanine and tryptophan to tyrosine and phenylalanine, did affect polymeric OAg production, especially in multiple-residue mutant constructs, indicating that they play a role in WecA function [5]. Similarly, cysteine is a polar, non-conservative mutation that could also increase the polarity of the region. Therefore, various cysteine substitutions in TMs IV and V appear to also diminish polymeric OAg production supporting the idea that multiple residues in this area play a role in WecA function.

3.4 PEGylation of WecA mutants

3.4.1 Optimization

PEGylation is a relatively recent approach for bacterial inner membrane proteins [89, 93, 95, 103, 104]. Therefore, it was necessary to optimize PEGylation of WecA before proceeding to the full characterization of the panel of mutants. First, we assessed the specificity of the label to make sure it could only detect exposed cysteine residues in the protein. The method employed is simplified to include labeling of EDTA-permeabilized cells and membrane preparations. The previously used and experimentally determined periplasmic control, G181C, was compared to the negative, cysteineless WecA encoded by pJL7 [20, 51]. Each control was labeled with increasing concentrations of PEG-Mal to demonstrate that PEGylation is functional and to determine the necessary amount of PEG-Mal that was required to obtain a band-shift. While 0.5 mM PEG-Mal was sufficient to visualize a detectable band-shift by Western blot, we chose to use 1 mM PEG-Mal in the labeling reactions to ensure there was

enough reagent available to PEGylate available cysteines in EDTA-permeabilized cells (Figure 16).

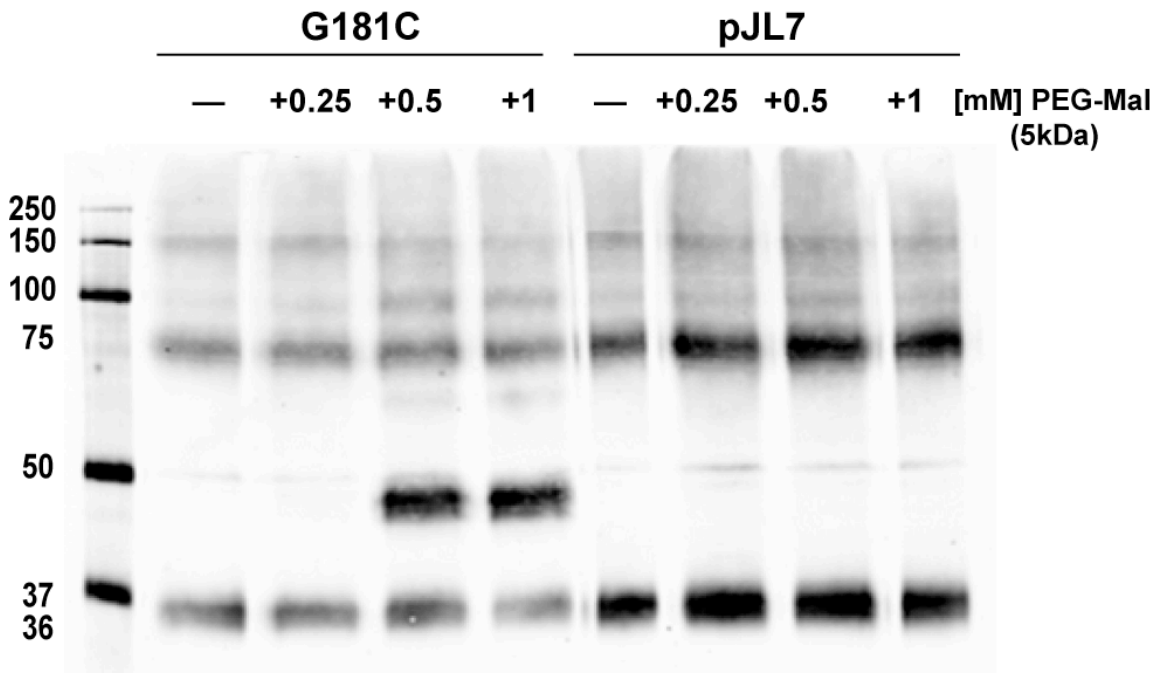


Figure 16. PEG-Mal labeling of cysteine-substituted WecA of *E. coli* in crude membrane preparations. The various cysteine-substituted mutants were disrupted by a cell disruptor set at 27 kPsi. The crude membrane fractions were isolated by ultracentrifugation and divided into four 20 μL aliquots. One aliquot was incubated with HEPES/NaCl buffer alone, whereas the remaining three aliquots were incubated with increasing concentrations of PEG-Mal (5 kDa). The reaction was quenched by addition of 45mM DTT, and a sample of each aliquot (5-10 μg of protein) was separated by 14% SDS-PAGE. Proteins were transferred to nitrocellulose membrane, and WecA was detected by anti-FLAG, with Alexa Fluor 680 anti-Mouse used as secondary antibody. The positions of the molecular mass markers are indicated on the left. G181C = periplasmic cysteine (should label with PEG-Mal); pJL7 = cysteine-less (should not label with PEG-Mal).

Second, we used additional replacements of known topology to evaluate if the PEGylation occurs as expected. The periplasmic control (G181C) should be labeled with the membrane-impermeable PEG-Mal in an EDTA-permeabilized cell, and in membrane preparations, and the PEG bulk will create an apparent band shift on a Western blot. The cytoplasmic control S362C should not be labeled with the membrane-impermeable PEG-Mal in EDTA-permeabilized cells, and not cause a band shift on a Western blot. Contrarily, S362C should be labeled with PEG-Mal in membrane preparations and cause a band shift as membrane preparations create mixed inverted vesicles allowing cytoplasmic exposed and bordering TM residues to react with PEG-Mal [20, 22, 23, 51, 92]. Finally, based on the topological prediction and experimental results employing SCAM, we chose F143C as the TM control [20]. F143C is predicted to be buried in a hydrophobic space and it was not accessible to biotinylation [20]; therefore, it should not react with PEG-Mal.

Both periplasmic (G181C) and cytoplasmic (S362C) controls reacted with PEG-Mal as expected. G181C showed an additional PEGylated double-band on Western blots in EDTA-permeabilized cells and in membrane preparations. S362C was not labeled in EDTA-permeabilized cells but showed a band-shift in membrane preparations. F143C was not labeled in EDTA-permeabilized cells, as expected. However, this residue consistently and repeatedly showed a double-band in membrane preparations (Figure 17). This suggests that F143C is partially exposed to the cytosol, or is at least facing outwards from the lipid bilayer, so that PEG-Mal can react with the substituted cysteine in the presence of water.

Despite that the expected "TM control" denoted a residue that was accessible from the cytosol (see below), the results with the periplasmic- and cytoplasmic-exposed residues confirmed their previously predicted location. Therefore, the optimization experiments suggested that PEGylation is an attractive alternative to biotin-SCAM, as it eliminates the need for purification.

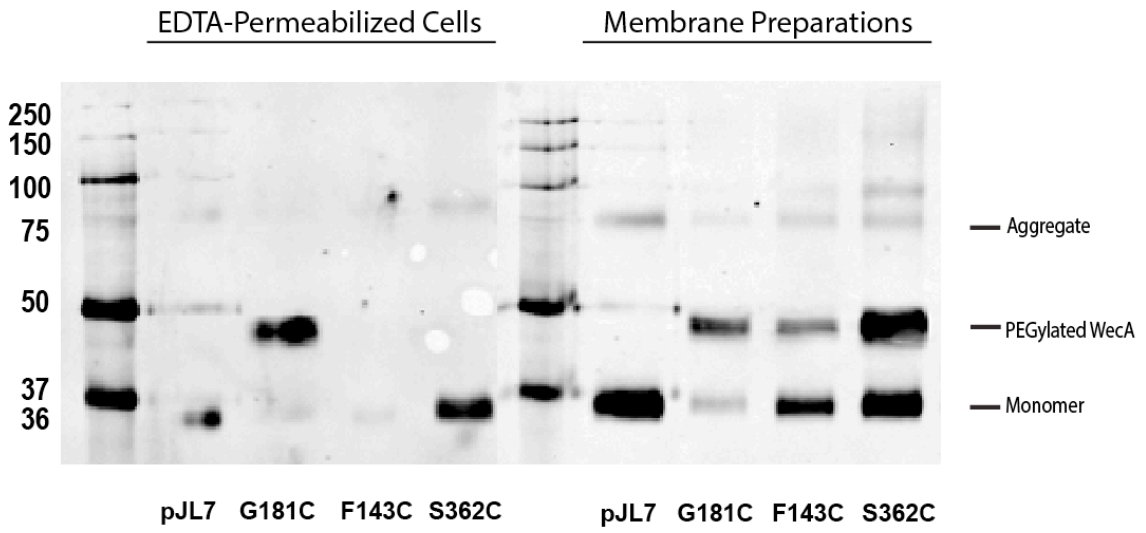


Figure 17. PEG-Mal labeling of cysteine-substituted WecA in *E. coli* controls in EDTA-permeabilized cells. The various cysteine-substituted mutants were grown until mid-exponential phase until protein expression was induced by arabinose. Cells were harvested and resuspended in 1mL of HEPES/$MgCl_2$ buffer and divided into aliquots of an OD of 5. One aliquot was incubated with incubated with 1mM PEG-Mal (5 kDa) and 5mM EDTA. The reaction was quenched by addition of 45mM DTT, then the membrane protein fractions were isolated by ultracentrifugation, and a sample of each aliquot (5-10µg of protein) was separated by 14% SDS-PAGE. The other aliquot was treated with 1mM PEG-Mal (5 kDa) after the membrane protein fractions were isolated, and then the reaction was quenched with 45mM DTT before a sample of each aliquot (5-10µg of protein) was separated by 14% SDS-PAGE. Proteins were transferred to nitrocellulose membrane, and WecA was detected by anti-FLAG, with Alexa Fluor 680 anti-Mouse used as secondary antibody. The positions of the molecular mass markers are indicated on the left, the condition on the top, aggregates, monomers and PEGylated WecA on the right, and PEGylated WecA mutant derivatives at the bottom. pJL7 = cysteine-less WecA (negative control); G181C = periplasmic cysteine; F143C = TM cysteine; S362C = cytoplasmic cysteine.

3.4.2 Cytoloop 1

In a previous study [19], the topological location of aspartic acid 35 differed experimentally from the *in silico* prediction. Biotin-SCAM suggested that this residue is located in the periplasm, in contrast to the location in cytoloop 1 predicted *in silico*. Also, replacement of aspartic acid 35 with cysteine or glycine resulted in reduced OAg production in *E. coli* MV501 and also reduced enzymatic activity *in vitro*, suggesting this residue plays a role in WecA function [20, 23]. Therefore, it was important to determine the location of Asp-35 as well defining the boundaries of cytoloop 1. For this purpose, several replacements were made in residues of the predicted cytoloop 1, which included: A24C, R25C, V27C, K30C, D35C, K41C, Q44C, G45C, and L46C (Figure 18). More specifically, A24C and R25C do not show labeling in intact cells or membrane preparations, suggesting they are within TM I. V27C does not show any labeling in intact cells but does in membrane preparations suggesting a cytoplasmic orientation.

Poor expression of Q44C and G45C may indicate that these residues are located at the border of TM II, which can cause the protein to be unstable and can make the PEGylation results inconclusive. Mutating these residues to alanine should restore the stability of the protein, indicating that they are important to WecA stability and are located at the TM border. K30C, D35C, K41C, and L46C all label in membrane preparations, but not intact cells, again suggesting a cytoplasmic orientation. Therefore, K30C, D35C and K41C are located in the cytoplasm, matching the predicted location. Q44C, G45C, and L46C are most likely found at the border of TM II, as predicted, although labeling of L46C suggests that it is exposed to the cytoplasm. This could be due

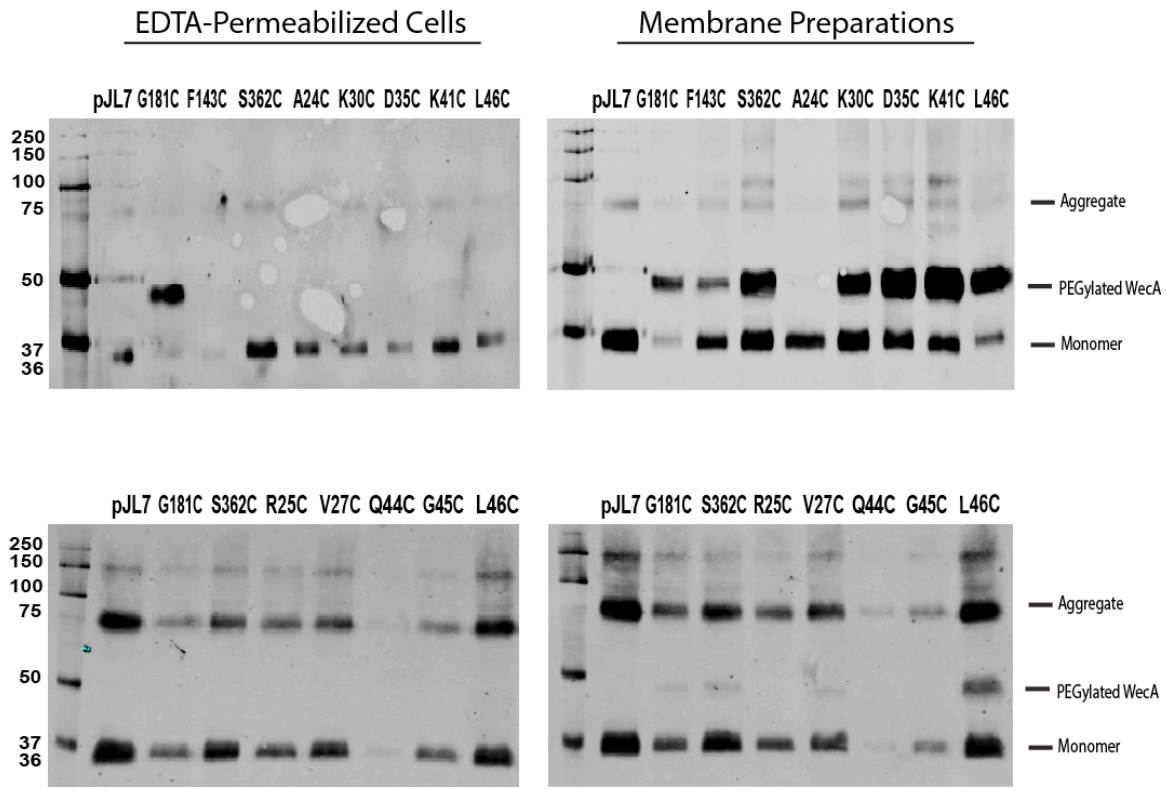


Figure 18. PEG-Mal labeling of cytoloop 1 of *E. coli* cysteine-substituted WecA in EDTA-permeabilized cells and crude membrane preparations. Samples were prepared and blotted as described above in figure 17. The positions of the molecular mass markers are indicated on the left, aggregates, monomers and PEGylated WecA on the right, and WecA derivatives on the top of each blot. pJL7 = cysteine-less WecA; G181C = periplasmic control; S362C = cytoplasmic control

to L46C being partially exposed to the cytosol, as it is located at the border of the TM and is therefore available to react with PEG-Mal [89]. These results demonstrate the specificity and sensitivity of the PEGylation method, and support the predicted cytoplasmic orientation of the D35 residue.

3.4.3 TMs I, II and III

Additional single-cysteine mutations were created throughout TMs I, II and III in the WecA protein, as these regions have not been thoroughly investigated. The mutations include: L4C, S7C, S12C, L19C, S54C, F61C, V66C, Y69C, A73C, L77C, and L83C. Based on the PEGylation results in EDTA-permeabilized cells and membrane preparations, all the residues are located in their predicted locations, except L4C which could potentially be included within the boundary of TM I based on the lack of band-shift pattern in EDTA-permeabilized cells (Figure 19). Interestingly, L4C and S7C show a faint band-shift in membrane preparations indicating some exposure to PEG-Mal. This is most likely due to their proximity to the N-terminus of WecA as membrane preparations create mixed inverted vesicles such that any residues close to the edge of the TM border can be partially exposed and therefore available to interact with PEG-Mal. This is also seen with V66C and A73C in membrane preparations while Y69C is labeled in EDTA-permeabilized cells supporting its predicted periplasmic orientation. Therefore, with the exception of residues located at the TM borders that may account for the band-shift patterns seen in membrane preparations, PEGylation supports the predicted locations of all the residues in TMs I, II and III. In addition, the results provide experimental refinement supporting the topological assignment of this region.

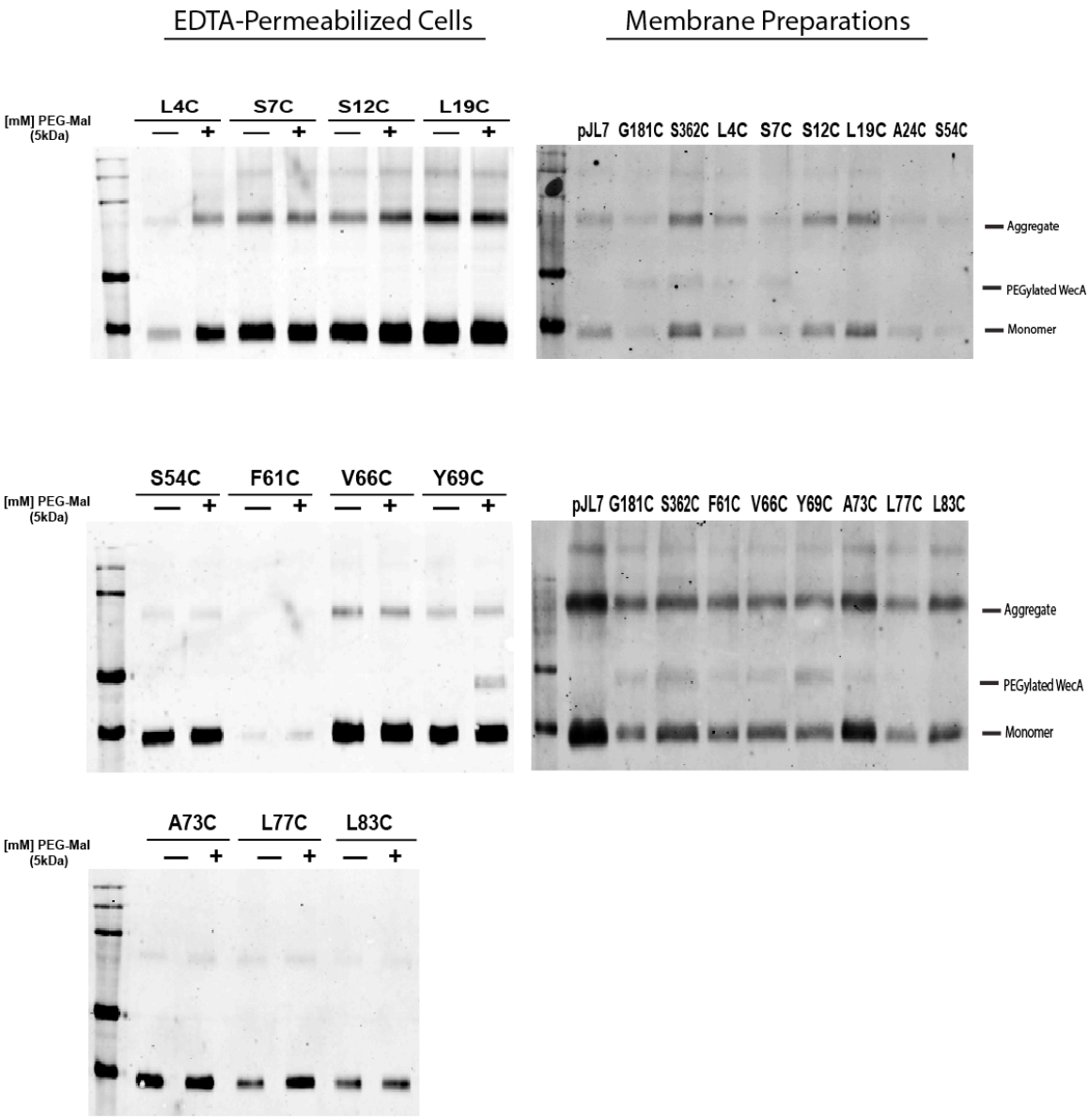


Figure 19. PEG-Mal labeling of TMs I, II, and III of *E. coli* cysteine-substituted WecA in EDTA-permeabilized cells and crude membrane preparations. Samples were prepared and blotted as described above in figure 17. The positions of the molecular mass markers are indicated on the left, aggregates, monomers and PEGylated WecA on the right, and WecA derivatives on the top of each blot. pJL7 = cysteine-less WecA; G181C = periplasmic control; S362C = cytoplasmic control

3.4.4 TMs IV and V

Currently, there is little known about the physical interaction between isoprenoids and their enzymes. The direct interaction between Und-P and WecA had not been demonstrated even though molecular modeling techniques have revealed specific interactions between hydrophobic residues in the PIRS peptides and the isoprenoids of other proteins. Residues located in TMs IV and V are of interest as they are in an arrangement that resembles a putative PIRS motif [102, 105, 106]. Both TMs are rich in aromatic residues, and we have previously speculated they may be partially exposed to the cytoplasm creating an interaction site ("pocket") with the α-isoprene unit of Und-P [5]. Previous experimental attempts show complementation defects and transferase assays as support for the hypothesis that Und-P interacts with WecA, however kinetic analyses remain inconclusive [5]. This is largely due to the fact that WecA cannot be purified to homogeneity in a large enough quantity to perform these functional analyses. Therefore, another objective of this study was to utilize the PEGylation protocol to support the hypothesis that Und-P interacts specifically with WecA via multiple hydrophobic interactions, as purification is not required. In fact, PEGylation of this region facilitates additional desired structural information of WecA. So, most of the residues mutated along TMs IV and V were originally predicted to be part of the putative PIRS motif [5] in addition to other surrounding residues in this region.

V107, V110, M112 and F114 label in membrane preparations but not in EDTA-permeabilized cells suggesting exposure to the cytosol. Q104 is PEGylated in membrane preparations, confirming it is on the edge of the cytoplasmic border of TM IV (Figure 20).

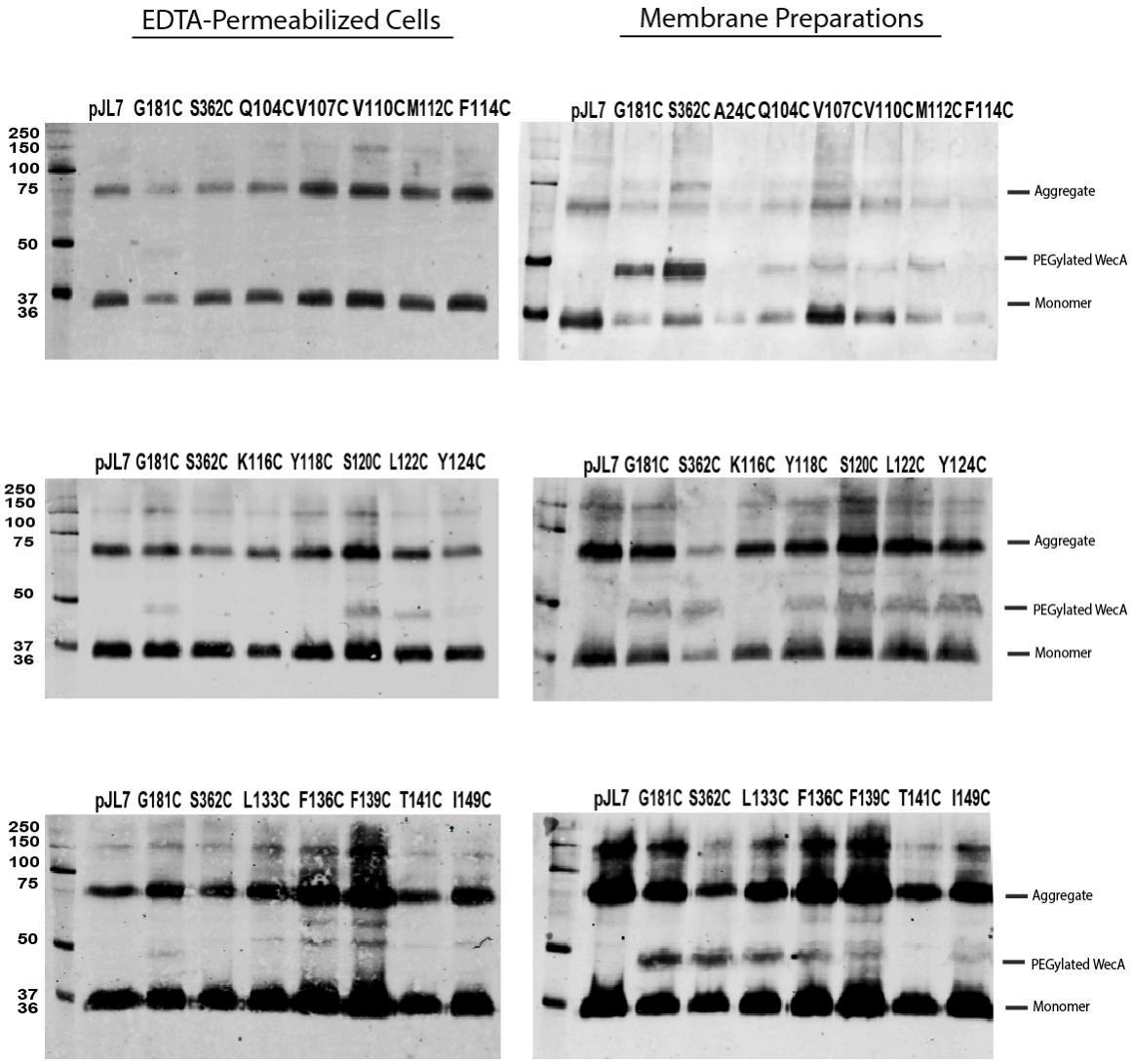


Figure 20. PEG-Mal labeling of TMs 4 and 5 (PIRS motif) of *E. coli* cysteine-substituted WecA in EDTA-permeabilized cells and crude membrane preparations. Samples were prepared and blotted as described above in figure 17. The positions of the molecular mass markers are indicated on the left, aggregates, monomers and PEGylated WecA on the right, and WecA derivatives on the top of each blot. pJL7 = cysteine-less WecA; G181C = periplasmic control; S362C = cytoplasmic control.

K116 does not label in either condition, confirming its predicted location in TM IV and most likely facing into the lipid bilayer. Next, Y118 does not label in EDTA-permeabilized cells but does show a band-shift in membrane preparations suggesting it is exposed and part of the PIRS motif. Next, S120, L122 and Y124 show a band-shift in both conditions suggesting that these residues are outside on TM IV and exposed to the periplasm (Figure 19). It is quite possible with the way WecA could be folded and tilted in the membrane that this part of TM IV is exposed.

The TM V residues L133, F136, F139 and I149 do not label in EDTA-permeabilized cells but do label in membrane preparations suggesting exposure from the cytoplasm. T141 does not show a band-shift in either condition confirming its predicted TM location and is therefore facing into the bilayer. F143 shows a band-shift in membrane preparations, and is therefore also exposed to the cytoplasm (as discussed above in section 3.4.1). This can be further supported by the fact that F143C does not complement OAg production [20, 92] and is therefore a critical residue for WecA function.

3.4.5 Cytoloop 4 (VFMGD motif)

Arginine 209 [20, 92] and the highly conserved VFMGD motif [51] include additional controversial regions that differed experimentally from the predicted topology. R209 was predicted to be in the cytosol while SCAM placed it in the periplasm. Contrarily, the VFMGD motif was predicted to be in TM VIII although it was thought to have a catalytic nucleophile based on modeling against the VFMGD motif in MraY [51, 72]. It is now partially reassigned to the cytosol as biotinylation of aspartic acid 217 provided indications of a cytosolic location. However, the other residues were

inaccessible to biotinylation, making their topological assignment difficult. Additional experiments indicate that the highly conserved VFMGD motif is likely a region in PNPT proteins that is involved in the binding and/or recognitions of the nucleotide moiety of the nucleoside phosphate precursor by contributing to the active site [51].

Using this information, residues R209C, Y211C, V213C, M215C, D217C, and S220C were assessed with the PEGylation protocol. None of these substituted cysteine residues label in EDTA-permeabilized cells, but they label in membrane preparations, suggesting the entire region is located in the cytoplasm (Figure 21). Based on the labeling pattern, S220C may be right at the edge of TM VIII as its band-shift is less intense compared to the other cytoplasmic residues but is still partially exposed to cytosol, allowing it to react with PEG-Mal [89]. The PEGylation results support the proposed cytoplasmic orientation of the VFMGD motif. This location is consistent with the proposed role of the region in the binding and/or recognition of nucleotide moiety as it would need to be exposed to the cytosol to interact with the nucleoside phosphate precursor. Additionally, if all the residues of this highly conserved region are located in the cytoplasm then the positively charged R209, two residues away, cannot be in the periplasm. Therefore, PEGylation of R209C supports the predicted cytosolic location.

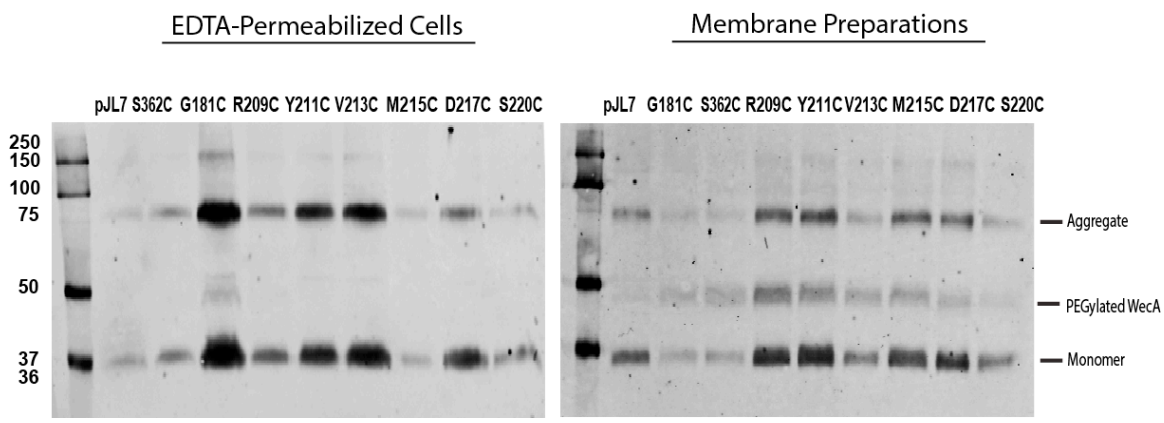


Figure 21. PEG-Mal labeling of cytoloop 4 of *E. coli* cysteine-substituted WecA in EDTA-permeabilized cells and crude membrane preparations. Samples were prepared and blotted as described above in figure 17. The positions of the molecular mass markers are indicated on the left, aggregates, monomers and PEGylated WecA on the right, and WecA derivatives on the top of each blot. pJL7 = cysteine-less WecA; G181C = periplasmic control; S362C = cytoplasmic control

Chapter 4

Discussion

Discussion

Elucidating the topology of WecA and uncovering critical functional residues in this protein provides a template to better understand the mechanism of action of PNPT family members. This study involved working with WecA in its native environment at the inner membrane. Currently, the structural characterization of PNPTs has been hampered by difficulties encountered with their overexpression and purification [18, 92]. Many previous topological methods miss important biological interactions (*in silico*), disrupt secondary structure of the protein (reporter fusions), and require purification (SCAM with biotinylation). PEGylation provides the advantage that it does not require protein purification, as it allows direct detection of labeled cysteine residues by a 10-25 kDa shift in the apparent mass of the protein in its native membrane location.

Aspartic acid 35 and arginine 209 are examples of WecA residues that differed experimentally from the predicted topology. The biotin-SCAM suggested that both residues were located in the periplasm, while the topology software predicted a cytosolic location [20, 92]. Replacement of aspartic acid 35 with cysteine resulted in loss of complementation (O7 LPS synthesis in *E. coli* MV501) [20, 92]. PEGylation results of this work support a cytoplasmic location of both D35 and R209 as these residues only labeled with PEG-Mal in membrane preparations, and not EDTA-permeabilized cells. We interpreted the differences between biotinylation-SCAM and PEGylation-SCAM as due to protein denaturation during purification.

Additional mutations were created along predicted cytoloop 1, including the borders of TMs I and II, to experimentally define the boundaries of the loop and the borders of the adjacent TMs. The borders of TM I can be updated to include L4, A24 and

R25 within the TM while V27 is located in the cytoplasm. This demonstrates the sensitivity and specificity of PEGylation as only one residue K26 separated the cytoplasmic border of TM I. At the other end of cytoloop 1 Q44, G45, and L46 are all labeled in membrane preparations, and not EDTA-permeabilized cells, indicating a cytoplasmic orientation and again refining the border to TM II.

Next, the highly conserved VFMGD motif was experimentally assessed with PEGylation as it is now partially located in the cytosol as previous biotinylation of aspartic acid 217 greatly changed the putative TM boundaries. The other residues were inaccessible to biotinylation, and may be in the TM, while the border on the other side of the VFMGD motif remains unrefined [51]. Again, using PEGylation, the entire region of substituted cysteines can be placed within cytoloop 4. Even S220 is labeled, albeit not as strongly, indicating that the cytoplasmic border of TM VIII can be updated. Moreover, all the substituted cysteines in the VFMGD motif and surrounding area tested with PEGylation provided conclusive results, unlike previous experimental attempts using biotinylation.

Initially, F143C was chosen as a control residue in a TM location because previous experimental evidence using SCAM with biotinylation suggested it was inaccessible. When labeling F143C with PEG-Mal in EDTA-permeabilized cells, there was no observable band-shift, as expected for a transmembrane location. However, F143C was consistently and reproducibly labeled in PEGylated membrane preparations. Previous research in our laboratory has suggested that F143 might be part of a region corresponding to a potential Und-P binding site [5], which spans TMs IV and V. These two TMs contains an unusual density of bulky aromatic residues, such as phenylalanine

and tryptophan, which may be important in interacting with the Und-P lipid substrate by creating a "pocket" accessible from the cytoplasm. Our results also suggest that these residues in both TMs would face each other rather than establishing interactions with membrane glycerophospholipids, which will make the residues inaccessible to PEG-Mal labeling.

After the experimental work in this study was completed, we submitted the WecA amino acid sequence to the iterative threading assembly refinement (I-TASSER) server for *ab initio* three-dimensional (3D) modeling. I-TASSER is an integrated web-based platform for fully automated protein structure and function prediction. It generates a 3D atomic model from multiple threading alignments and iterative structural assembly simulations using an amino acid sequence without the common biochemical limits set by human intervention. Structurally matching the 3D model with other known proteins allows the function of the protein of interest to be inferred. Full-length secondary and tertiary structure predictions and functional annotations on ligand-binding sites are some of the usual outputs from a typical server. Moreover, I-TASSER is consistently ranked as the best method for computational structural biology [107, 108].

Comparing the PEGylation results with the 3D model demonstrated a remarkable agreement between experimental data and the *ab initio* structural prediction. The experimentally assigned location of nearly every residue that was mutated to cysteine and PEGylated agreed with the structural model. Moreover, the structural model supported the hypothesis that many of the residues located in TMs IV and V, especially the phenylalanines form a "cavity" opened to the cytoplasmic space, which could accommodate the lipid carrier. Additionally, the three residues S120, L122, and Y124 at

the border of TM IV that show a band-shift in EDTA-permeabilized cells are supported by the model to be exposed to the periplasm in loop 2 (see Appendix A). This recent automated 3D model is well supported by the current experimental data, suggesting that PEGylation is sensitive, specific and accurate method to elucidate the topology of a membrane protein like WecA.

In summary, the PEGylation approach employed here has allowed us to characterize and refine the topology of WecA. Currently, four other integral membrane proteins are being topologically mapped in our laboratory using the PEGylation protocol including *E. coli* WaaL, *Pseudomonas aeruginosa* WaaL, *E. coli* WcaJ, and *Burkholderia cenocepacia* ArnT.

References

1. Nikaido, H., *Prevention of drug access to bacterial targets: permeability barriers and active efflux.* Science, 1994. **264**(5157): p. 382-8.

2. Nikaido, H. and M. Vaara, *Molecular basis of bacterial outer membrane permeability.* Microbiol Rev, 1985. **49**(1): p. 1-32.

3. Martin, E.L. and R.A. MacLeod, *Isolation and chemical composition of the cytoplasmic membrane of a gram-negative bacterium.* J Bacteriol, 1971. **105**(3): p. 1160-7.

4. Vollmer, W., D. Blanot, and M.A. de Pedro, *Peptidoglycan structure and architecture.* FEMS Microbiol Rev, 2008. **32**(2): p. 149-67.

5. Haggerty, E., *Characterization of a potential undecaprenyl phosphate binding site on the Escherichia coli O-antigen initiation enzyme, WecA*, in *Microbiology and Immunology*. 2008, *The* University *of* Western Ontario: London, ON, Canada.

6. Kuhn, H.M., U. Meier-Dieter, and H. Mayer, *ECA, the enterobacterial common antigen.* FEMS Microbiol Rev, 1988. **4**(3): p. 195-222.

7. Makela, P.H. and H. Mayer, *Enterobacterial common antigen.* Bacteriol Rev, 1976. **40**(3): p. 591-632.

8. Mannel, D. and H. Mayer, *Isolation and chemical characterization of the enterobacterial common antigen.* Eur J Biochem, 1978. **86**(2): p. 361-70.

9. Whitfield, C., *Biosynthesis and assembly of capsular polysaccharides in Escherichia coli.* Annu Rev Biochem, 2006. **75**: p. 39-68.

10. Whitfield, C. and I.S. Roberts, *Structure, assembly and regulation of expression of capsules in Escherichia coli.* Mol Microbiol, 1999. **31**(5): p. 1307-19.

11. Hornef, M.W., et al., *Bacterial strategies for overcoming host innate and adaptive immune responses.* Nat Immunol, 2002. **3**(11): p. 1033-40.

12. Joiner, K.A., *Complement evasion by bacteria and parasites.* Annu Rev Microbiol, 1988. **42**: p. 201-30.

13. Raetz, C.R. and C. Whitfield, *Lipopolysaccharide endotoxins.* Annu Rev Biochem, 2002. **71**: p. 635-700.

14. Holst, O., et al., *Biochemistry and cell biology of bacterial endotoxins.* FEMS Immunol Med Microbiol, 1996. **16**(2): p. 83-104.

15. Orskov, I., et al., *Serology, chemistry, and genetics of O and K antigens of Escherichia coli.* Bacteriol Rev, 1977. **41**(3): p. 667-710.

16. Samuel, G. and P. Reeves, *Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly.* Carbohydr Res, 2003. **338**(23): p. 2503-19.

17. Valvano, M.A., S. E. Furlong and K. B. Patel, *Genetics, Biosynthesis and Assembly of O-antigen*, in *Bacterial Lipopolysaccharides*, Y.A.a.M.A.V. Knirel, Editor. 2011, Springer/Wien: New York. p. 275-310.

18. Al-Dabbagh, B., D. Mengin-Lecreulx, and A. Bouhss, *Purification and characterization of the bacterial UDP-GlcNAc:undecaprenyl-phosphate GlcNAc-1-phosphate transferase WecA.* J Bacteriol, 2008. **190**(21): p. 7141-6.

19. Hug, I. and M.F. Feldman, *Analogies and homologies in lipopolysaccharide and glycoprotein biosynthesis in bacteria.* Glycobiology. **21**(2): p. 138-51.

20. Lehrer, J., et al., *Functional characterization and membrane topology of Escherichia coli WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide.* J Bacteriol, 2007. **189**(7): p. 2618-28.

21. Valvano, M.A., *Common themes in glycoconjugate assembly using the biogenesis of O-antigen lipopolysaccharide as a model system.* Biochemistry (Mosc). **76**(7): p. 729-35.

22. Amer, A.O. and M.A. Valvano, *Conserved amino acid residues found in a predicted cytosolic domain of the lipopolysaccharide biosynthetic protein WecA are implicated in the recognition of UDP-N-acetylglucosamine.* Microbiology, 2001. **147**(Pt 11): p. 3015-25.

23. Amer, A.O. and M.A. Valvano, *Conserved aspartic acids are essential for the enzymic activity of the WecA protein initiating the biosynthesis of O-specific lipopolysaccharide and enterobacterial common antigen in Escherichia coli.* Microbiology, 2002. **148**(Pt 2): p. 571-82.

24. Batchelor, R.A., et al., *Regulation by a novel protein of the bimodal distribution of lipopolysaccharide in the outer membrane of Escherichia coli.* J Bacteriol, 1991. **173**(18): p. 5699-704.

25. Feldman, M.F., et al., *The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in Escherichia coli O antigen assembly is independent of the chemical structure of the O repeat.* J Biol Chem, 1999. **274**(49): p. 35129-38.

26. Klena, J.D., R.S. Ashford, 2nd, and C.A. Schnaitman, *Role of Escherichia coli K-12 rfa genes and the rfp gene of Shigella dysenteriae 1 in generation of lipopolysaccharide core heterogeneity and attachment of O antigen.* J Bacteriol, 1992. **174**(22): p. 7297-307.

27. Klena, J.D. and C.A. Schnaitman, *Function of the rfb gene cluster and the rfe gene in the synthesis of O antigen by Shigella dysenteriae 1.* Mol Microbiol, 1993. **9**(2): p. 393-402.

28. Liu, D., R.A. Cole, and P.R. Reeves, *An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase.* J Bacteriol, 1996. **178**(7): p. 2102-7.

29. Macpherson, D.F., P.A. Manning, and R. Morona, *Genetic analysis of the rfbX gene of Shigella flexneri.* Gene, 1995. **155**(1): p. 9-17.

30. Marolda, C.L., et al., *Functional analysis of predicted coiled-coil regions in the Escherichia coli K-12 O-antigen polysaccharide chain length determinant Wzz.* J Bacteriol, 2008. **190**(6): p. 2128-37.

31. Marolda, C.L., J. Vicarioli, and M.A. Valvano, *Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit.* Microbiology, 2004. **150**(Pt 12): p. 4095-105.

32. Morona, R., L. Van Den Bosch, and C. Daniels, *Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions.* Microbiology, 2000. **146 (Pt 1)**: p. 1-4.

33. Schnaitman, C.A. and J.D. Klena, *Genetics of lipopolysaccharide biosynthesis in enteric bacteria.* Microbiol Rev, 1993. **57**(3): p. 655-82.

34. Tocilj, A., et al., *Bacterial polysaccharide co-polymerases share a common framework for control of polymer length.* Nat Struct Mol Biol, 2008. **15**(2): p. 130-8.

35. Whitfield, C., P.A. Amor, and R. Koplin, *Modulation of the surface architecture of gram-negative bacteria by the action of surface polymer:lipid A-core ligase and by determinants of polymer chain length.* Mol Microbiol, 1997. **23**(4): p. 629-38.

36. Valvano, M.A., *Export of O-specific lipopolysaccharide.* Front Biosci, 2003. **8**: p. s452-71.

37. Reizer, J., A. Reizer, and M.H. Saier, Jr., *A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates.* Protein Sci, 1992. **1**(10): p. 1326-32.

38. Jiang, X.M., et al., *Structure and sequence of the rfb (O antigen) gene cluster of Salmonella serovar typhimurium (strain LT2).* Mol Microbiol, 1991. **5**(3): p. 695-713.

39. Patel, K.B., S.E. Furlong, and M.A. Valvano, *Functional analysis of the C-terminal domain of the WbaP protein that mediates initiation of O antigen synthesis in Salmonella enterica.* Glycobiology. **20**(11): p. 1389-401.

40. Saldias, M.S., et al., *Distinct functional domains of the Salmonella enterica WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit.* Microbiology, 2008. **154**(Pt 2): p. 440-53.

41. Wang, L., D. Liu, and P.R. Reeves, *C-terminal half of Salmonella enterica WbaP (RfbP) is the galactosyl-1-phosphate transferase domain catalyzing the first step of O-antigen synthesis.* J Bacteriol, 1996. **178**(9): p. 2598-604.

42. Patel, K.B., et al., *The C-terminal domain of the Salmonella enterica WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate.* Glycobiology. **22**(1): p. 116-22.

43. Stevenson, G., et al., *Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid.* J Bacteriol, 1996. **178**(16): p. 4885-93.

44. Drummelsmith, J. and C. Whitfield, *Gene products required for surface expression of the capsular form of the group 1 K antigen in Escherichia coli (O9a:K30).* Mol Microbiol, 1999. **31**(5): p. 1321-32.

45. Patel, K.B., et al., *Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of Escherichia coli and Caulobacter crescentus.* J Bacteriol. **194**(10): p. 2646-57.

46. Cartee, R.T., et al., *CpsE from type 2 Streptococcus pneumoniae catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation.* J Bacteriol, 2005. **187**(21): p. 7425-33.

47. Xayarath, B. and J. Yother, *Mutations blocking side chain assembly, polymerization, or transport of a Wzy-dependent Streptococcus pneumoniae capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall.* J Bacteriol, 2007. **189**(9): p. 3369-81.

48. Steiner, K., et al., *Functional characterization of the initiation enzyme of S-layer glycoprotein glycan biosynthesis in Geobacillus stearothermophilus NRS 2004/3a.* J Bacteriol, 2007. **189**(7): p. 2590-8.

49. Rush, J.S., P.D. Rick, and C.J. Waechter, *Polyisoprenyl phosphate specificity of UDP-GlcNAc:undecaprenyl phosphate N-acetylglucosaminyl 1-P transferase from E.coli.* Glycobiology, 1997. **7**(2): p. 315-22.

50. Price, N.P. and F.A. Momany, *Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases.* Glycobiology, 2005. **15**(9): p. 29R-42R.

51. Furlong, S.E. and M.A. Valvano, *Characterization of the highly conserved VFMGD motif in a bacterial polyisoprenyl-phosphate N-acetylaminosugar-1-phosphate transferase.* Protein Sci. **21**(9): p. 1366-75.

52. Hubbard, S.C. and R.J. Ivatt, *Synthesis and processing of asparagine-linked oligosaccharides.* Annu Rev Biochem, 1981. **50**: p. 555-83.

53. Lehrman, M.A., *A family of UDP-GlcNAc/MurNAc: polyisoprenol-P GlcNAc/MurNAc-1-P transferases.* Glycobiology, 1994. **4**(6): p. 768-71.

54. Clarke, B.R., et al., *Role of Rfe and RfbF in the initiation of biosynthesis of D-galactan I, the lipopolysaccharide O antigen from Klebsiella pneumoniae serotype O1.* J Bacteriol, 1995. **177**(19): p. 5411-8.

55. Yao, Z. and M.A. Valvano, *Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (rfb) of Escherichia coli K-12 W3110: identification of genes that confer group 6 specificity to Shigella flexneri serotypes Y and 4a.* J Bacteriol, 1994. **176**(13): p. 4133-43.

56. Zhang, L., et al., *Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of Yersinia enterocolitica serotype O:8.* Mol Microbiol, 1997. **23**(1): p. 63-76.

57. Alexander, D.C. and M.A. Valvano, *Role of the rfe gene in the biosynthesis of the Escherichia coli O7-specific lipopolysaccharide and other O-specific polysaccharides containing N-acetylglucosamine.* J Bacteriol, 1994. **176**(22): p. 7079-84.

58. Schmidt, G., H. Mayer, and P.H. Makela, *Presence of rfe genes in Escherichia coli: their participation in biosynthesis of O antigen and enterobacterial common antigen.* J Bacteriol, 1976. **127**(2): p. 755-62.

59. Amor, P.A. and C. Whitfield, *Molecular and functional analysis of genes required for expression of group IB K antigens in Escherichia coli: characterization of the his-region containing gene clusters for multiple cell-surface polysaccharides.* Mol Microbiol, 1997. **26**(1): p. 145-61.

60. Belanger, M., L.L. Burrows, and J.S. Lam, *Functional analysis of genes responsible for the synthesis of the B-band O antigen of Pseudomonas aeruginosa serotype O6 lipopolysaccharide.* Microbiology, 1999. **145 (Pt 12)**: p. 3505-21.

61. Ikeda, M., et al., *The Escherichia coli mraY gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-pentapeptide transferase.* J Bacteriol, 1991. **173**(3): p. 1021-6.

62. Rocchetta, H.L., et al., *Three rhamnosyltransferases responsible for assembly of the A-band D-rhamnan polysaccharide in Pseudomonas aeruginosa: a fourth transferase, WbpL, is required for the initiation of both A-band and B-band lipopolysaccharide synthesis.* Mol Microbiol, 1998. **28**(6): p. 1103-19.

63. Bugg, T.D. and P.E. Brandish, *From peptidoglycan to glycoproteins: common features of lipid-linked oligosaccharide biosynthesis.* FEMS Microbiol Lett, 1994. **119**(3): p. 255-62.

64. Dal Nogare, A.R., N. Dan, and M.A. Lehrman, *Conserved sequences in enzymes of the UDP-GlcNAc/MurNAc family are essential in hamster UDP-GlcNAc:dolichol-P GlcNAc-1-P transferase.* Glycobiology, 1998. **8**(6): p. 625-32.

65. Anderson, M.S., S.S. Eveland, and N.P. Price, *Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family.* FEMS Microbiol Lett, 2000. **191**(2): p. 169-75.

66. Zhu, X.Y. and M.A. Lehrman, *Cloning, sequence, and expression of a cDNA encoding hamster UDP-GlcNAc:dolichol phosphate N-acetylglucosamine-1-phosphate transferase.* J Biol Chem, 1990. **265**(24): p. 14250-5.

67. Bouhss, A., et al., *Topological analysis of the MraY protein catalysing the first membrane step of peptidoglycan synthesis.* Mol Microbiol, 1999. **34**(3): p. 576-85.

68. Eddy, S.R., *Hidden Markov models.* Curr Opin Struct Biol, 1996. **6**(3): p. 361-5.

69. Amer, A.O. and M.A. Valvano, *The N-terminal region of the Escherichia coli WecA (Rfe) protein, containing three predicted transmembrane helices, is required for function but not for membrane insertion.* J Bacteriol, 2000. **182**(2): p. 498-503.

70. Sekine, S., et al., *Crucial role of the high-loop lysine for the catalytic activity of arginyl-tRNA synthetase.* J Biol Chem, 2001. **276**(6): p. 3723-6.

71. Venkatachalam, K.V., et al., *Site-selected mutagenesis of a conserved nucleotide binding HXGH motif located in the ATP sulfurylase domain of human bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase.* J Biol Chem, 1999. **274**(5): p. 2601-4.

72. Lloyd, A.J., et al., *Phospho-N-acetyl-muramyl-pentapeptide translocase from Escherichia coli: catalytic role of conserved aspartic acid residues.* J Bacteriol, 2004. **186**(6): p. 1747-57.

73. Liang, J., et al., *Computational studies of membrane proteins: models and predictions for biological understanding.* Biochim Biophys Acta. **1818**(4): p. 927-41.

74. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein.* J Mol Biol, 1982. **157**(1): p. 105-32.

75. Cymer, F., A. Veerappan, and D. Schneider, *Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties.* Biochim Biophys Acta. **1818**(4): p. 963-73.

76. Simon, I., A. Fiser, and G.E. Tusnady, *Predicting protein conformation by statistical methods.* Biochim Biophys Acta, 2001. **1549**(2): p. 123-36.

77. Ehrmann, M., D. Boyd, and J. Beckwith, *Genetic analysis of membrane protein topology by a sandwich gene fusion approach.* Proc Natl Acad Sci U S A, 1990. **87**(19): p. 7574-8.

78. Haardt, M. and E. Bremer, *Use of phoA and lacZ fusions to study the membrane topology of ProW, a component of the osmoregulated ProU transport system of Escherichia coli.* J Bacteriol, 1996. **178**(18): p. 5370-81.

79. Manoil, C., *Analysis of protein localization by use of gene fusions with complementary properties.* J Bacteriol, 1990. **172**(2): p. 1035-42.

80. Silhavy, T.J. and J.R. Beckwith, *Uses of lac fusions for the study of biological problems.* Microbiol Rev, 1985. **49**(4): p. 398-418.

81. Bogdanov, M., P.N. Heacock, and W. Dowhan, *A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition.* Embo J, 2002. **21**(9): p. 2107-16.

82. Bogdanov, M., et al., *Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis.* Methods, 2005. **36**(2): p. 148-71.

83. Cao, W. and L.H. Matherly, *Analysis of the membrane topology for transmembrane domains 7-12 of the human reduced folate carrier by scanning cysteine accessibility methods.* Biochem J, 2004. **378**(Pt 1): p. 201-6.

84. Grunewald, M., D. Menaker, and B.I. Kanner, *Cysteine-scanning mutagenesis reveals a conformationally sensitive reentrant pore-loop in the glutamate transporter GLT-1.* J Biol Chem, 2002. **277**(29): p. 26074-80.

85. Loo, T.W. and D.M. Clarke, *Membrane topology of a cysteine-less mutant of human P-glycoprotein.* J Biol Chem, 1995. **270**(2): p. 843-8.

86. Sato, Y., et al., *Analysis of transmembrane domain 2 of rat serotonin transporter by cysteine scanning mutagenesis.* J Biol Chem, 2004. **279**(22): p. 22926-33.

87. Wang, X., M. Bogdanov, and W. Dowhan, *Topology of polytopic membrane protein subdomains is dictated by membrane phospholipid composition.* Embo J, 2002. **21**(21): p. 5673-81.

88. Culham, D.E., et al., *Creation of a fully functional cysteine-less variant of osmosensor and proton-osmoprotectant symporter ProP from Escherichia coli and its application to assess the transporter's membrane orientation.* Biochemistry, 2003. **42**(40): p. 11815-23.

89. Lu, J. and C. Deutsch, *Pegylation: a method for assessing topological accessibilities in Kv1.3.* Biochemistry, 2001. **40**(44): p. 13288-301.

90. Falke, J.J., et al., *Structure of a bacterial sensory receptor. A site-directed sulfhydryl study.* J Biol Chem, 1988. **263**(29): p. 14850-8.

91. Kumarevel, T.S., M.M. Gromiha, and M.N. Ponnuswamy, *Distribution of amino acid residues and residue-residue contacts in molecular chaperones.* Prep Biochem Biotechnol, 2001. **31**(2): p. 163-83.

92. Lehrer, J., *Topological mapping of the Escherichia coli WecA protein by site-directed sulfhydryl labeling*, in *Microbiology and Immunology*. 2005, *The* University *of* Western Ontario: London, ON, Canada.

93. Koch, S., et al., *Escherichia coli TatA and TatB proteins have N-out, C-in topology in intact cells.* J Biol Chem. **287**(18): p. 14420-31.

94. Leive, L., *Studies on the permeability change produced in coliform bacteria by ethylenediaminetetraacetate.* J Biol Chem, 1968. **243**(9): p. 2373-80.

95. Bauer, J., et al., *Topology and accessibility of the transmembrane helices and the sensory site in the bifunctional transporter DcuB of Escherichia coli.* Biochemistry. **50**(26): p. 5925-38.

96. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.

97. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning : a laboratory manual*. 2nd ed. 1989, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

98. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

99. Hitchcock, P.J. and T.M. Brown, *Morphological heterogeneity among Salmonella lipopolysaccharide chemotypes in silver-stained polyacrylamide gels.* J Bacteriol, 1983. **154**(1): p. 269-77.

100. Marolda, C.L., et al., *Genetic analysis of the O7-polysaccharide biosynthesis region from the Escherichia coli O7:K1 strain VW187.* J Bacteriol, 1990. **172**(7): p. 3590-9.

101. Tsai, C.M. and C.E. Frasch, *A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels.* Anal Biochem, 1982. **119**(1): p. 115-9.

102. Troy, F.A., 2nd, *Polysialylation: from bacteria to brains.* Glycobiology, 1992. **2**(1): p. 5-23.

103. Toei, M., S. Toei, and M. Forgac, *Definition of membrane topology and identification of residues important for transport in subunit a of the vacuolar ATPase.* J Biol Chem. **286**(40): p. 35176-86.

104. Wang, Y., M. Toei, and M. Forgac, *Analysis of the membrane topology of transmembrane segments in the C-terminal hydrophobic domain of the yeast vacuolar ATPase subunit a (Vph1p) by chemical modification.* J Biol Chem, 2008. **283**(30): p. 20696-702.

105. Zhou, G.P. and F.A. Troy, 2nd, *Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions.* Glycobiology, 2003. **13**(2): p. 51-71.

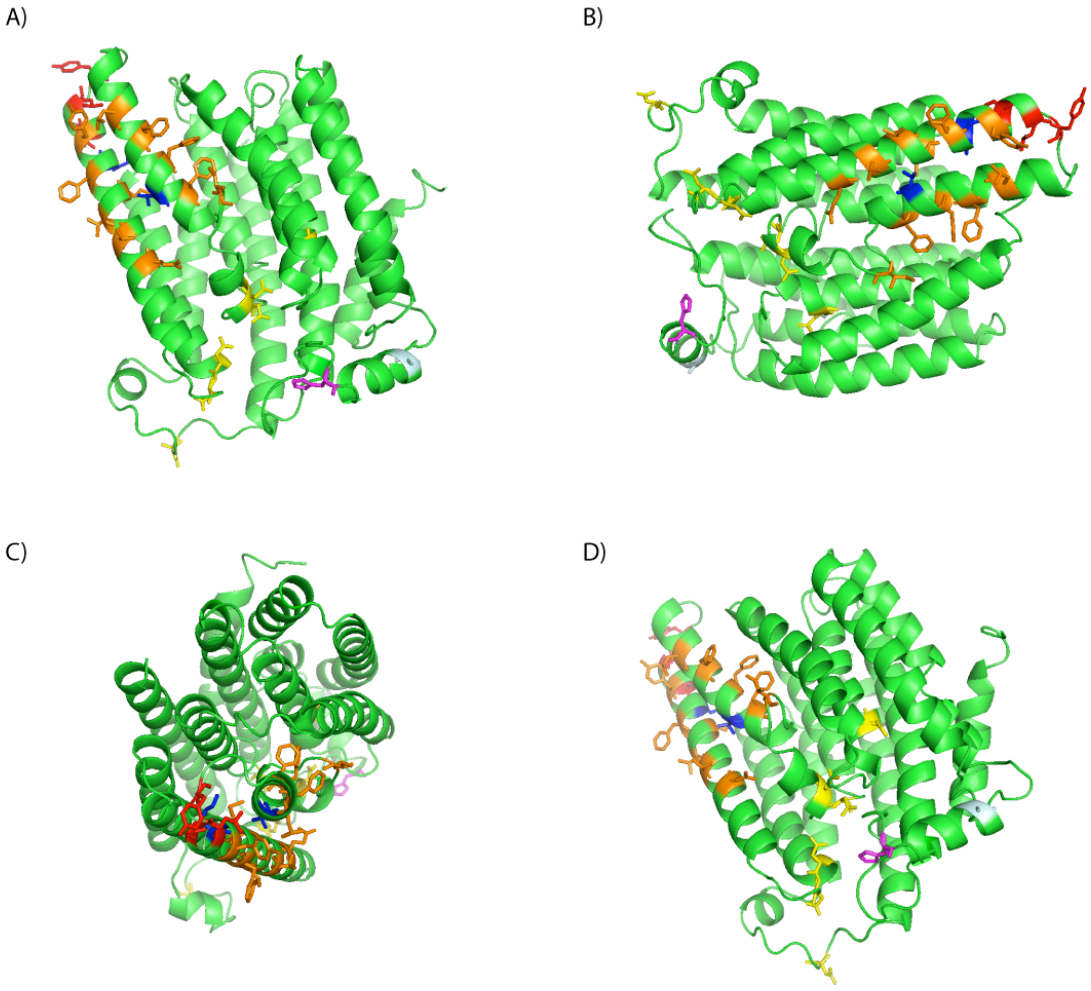
106. Zhou, G.P. and F.A. Troy, 2nd, *NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure.* Glycobiology, 2005. **15**(4): p. 347-59.

107. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nat Protoc. **5**(4): p. 725-38.

108. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.

Appendices

Appendix A



***Ab initio* 3D model of WecA.** The 3D model was derived using the I-TASSER server. A) View from the side. B) View through the bilayer. C) View from the periplasm. D) Tilted view from the side. The residues located along TMs IV and V that label in membrane preparations (suggesting exposure to the cytosol) are highlighted in orange. The residues located along TMs IV and V that do not label in membrane preparations (suggesting they are facing into the bilayer) are highlighted in blue. The residues S120, L122, and Y124 located along TMs IV and V that label in EDTA-permeabilized cells (suggesting exposure to the periplasm) are highlighted in red. Other critical aspartic acids D35, D90, D91, D156, and D159 are highlighted in yellow while H278 is highlighted in magenta and A286 is highlighted in pale blue.

VITA

Name: Stéphanie Lynn Lamothe

Post-secondary Education and Degrees:

The University *of* Western Ontario
London, Ontario, Canada
2005-2009 B.HSc

Nipissing University
North Bay, Ontario, Canada
2010-2011 Non-Degree Studies

The University *of* Western Ontario
London, Ontario, Canada
2011-2013 M.Sc

Honours and Awards:

Western Graduate Research Scholarship
The University *of* Western Ontario
2011-2013

Schulich School of Medicine and Dentistry
Graduate Scholarship
The University *of* Western Ontario
2011-2013